

REVIEW ARTICLE

The Mathematical Theory of Causation¹

Daniel M. Hausman

- 1 *An SGS primer*
 - 2 *What is so provocative?*
 - 3 *Philosophical criticisms*
 - 3.1 *The causal interpretation of induced probabilistic dependencies*
 - 3.2 *The interpretation of causation and plausibility of the causal Markov*
 - 3.3 *Probabilistic dependency, causation, and manipulation*
 - 4 *Conclusion*
-

Despite its unnecessarily melodramatic title, this is a superb collection. Most of the papers were originally presented at a raucous conference held at Notre Dame in October 1993. Others were added later to provide background and to expand arguments sketched at the conference. The heated disagreements at the conference have not been resolved, but the issues that divided the participants are now better articulated than they were five years ago, and most of them can be readily understood by philosophers, statisticians, and social scientists who are not experts in the technical literature concerning the possibility of making causal inferences from non-experimental data. Although all of these audiences have a good deal to learn from this volume, this review will address mainly philosophical issues.

The conference and volume both focus on the work of Peter Spirtes, Clark Glymour, and Richard Scheines (SGS from now on), particularly their 1993 book, *Causation, Prediction, and Search*, but readers should not skip the essays in this volume that do not focus on the work of SGS, several of which are gems. Clifford Clogg and Adamantios Haritou point out that one cannot safely infer that z is a cause of y from the fact that a model that regresses y on a set of variables including z fits the data better than one that regresses y on a set of variables that does not include z . Nancy Cartwright argues for the

¹ Review of Vaughn R. McKim and Stephen P. Turner (eds) [1997]: *Causality in Crisis? Statistical Methods and the Search for Causal Knowledge in the Social Sciences*, Notre Dame: Notre Dame University Press.

provocative thesis that *every* regularity arises from a (causal) structure—that is, ‘a fixed (enough) arrangement of components with stable (enough) capacities’ (p. 343). In her view, this thesis complicates the task of causal inference, because inference relies on regularities; and regularities provide information only about particular arrangements of components. Causal inferences will thus break down when causal structures break down, and variables that influence causal structure cannot be treated as ordinary causes and effects. To illustrate what she means, consider an ordinary dimmer switch that rotates to adjust the brightness of a light and clicks in and out to turn the current on or off. Cartwright argues that rotating the switch and turning it off and on should not be regarded simply as two causes of the illumination of the chandelier, because the ‘machine’ relating the rotation to the illumination is only connected when the switch is turned on (p. 350f). Because causal relations hold only relative to the connection and persistence of particular structures or machines, causal inferences are necessarily limited and fragile.

In essays that have still less direct connection to the work of SGS, Stephen Turner and Mary Morgan provide fascinating histories of earlier methods that were used to draw causal inferences from statistical data. Their essays are of interest both to those who are interested in the history of causal inference and to those who are concerned with contemporary methods. Mary Morgan examines what was gained—and lost—when theorists shifted to making causal inferences on the basis of information concerning correlations rather than relying on graphs. Turner tells the cautionary tale of William Farr who found that cholera cases during the epidemic of 1849 were inversely proportional to elevation and who concluded that the disease was carried by miasmata in the air. The most powerful statistical relationships were in this instance badly misleading. Only by considering a variety of evidence bearing on the *mechanism* of transmission was John Snow able to establish that the disease was water-borne (pp. 36–41).

The main concern of *Causality in Crisis?* is, however, the work of SGS. This is expounded in essays by Richard Scheines and Clark Glymour, examined and criticized in essays by David Freedman, Paul Humphreys, and James Woodward, and defended in replies by Spirtes and Scheines and by Glymour. The exchanges are sharp and not always friendly, and the issues are serious. In what follows, I shall first sketch a fragment of SGS’s work and explain why it has disturbed their critics and then I shall turn to three philosophical criticisms of their work. A warning: the issues I shall be discussing are complicated, and in a relatively brief review essay I cannot do them justice.

1 An SGS primer

This section provides what I hope is a useful caricature of SGS’s work, which

leaves out many central features and inevitably loses much of the ‘flavour’ of what they are up to. For example, it will not mention causal graphs, which SGS use to represent causal relations. Those who want a more balanced, though still brief and accessible introduction, should read Richard Scheines’ essay in this volume.

Suppose one is concerned with causal relations among variables, such as wealth and delinquency, or rainfall and crop yields. Second, suppose that the causal relationships one is concerned with are all asymmetric and that there are no cycles. Third, suppose that the set of variables V is ‘causally sufficient’ in roughly the sense that all variables related to members of V only as common causes are included in V . Fourth, suppose each variable has its own probabilistically independent source of variation, which is not itself included in V . With the help of two axioms, SGS can then relate the causal relations among members of V to features of the joint probability distribution of the members of V . These two axioms are:

Causal Markov condition: For all x, y in V , if x does not cause y , then x and y are probabilistically independent conditional on the subset of V consisting of all the direct causes of x .

Faithfulness condition: For all x, y in V , if x and y are probabilistically independent conditional on the subset of V consisting of all the direct causes of x , then x does not cause y .

x is a ‘direct’ cause of y if x exerts a causal influence on y that is not mediated by any of the other variables in V . Direct causes need not be ‘immediate’ in any absolute or spatio-temporal sense. The causal Markov and faithfulness conditions jointly provide a truth condition for ‘ x causes y ’ provided that the set of variables V is causally sufficient. x causes y if and only if the probabilistic dependency between them is not screened off by the set of all the direct causes of x . This truth condition obviously does not constitute an analysis of causation, since the term ‘causes’ appears on both sides of the biconditional, and SGS claim that their work is compatible with a variety of theories of causation.

The causal Markov condition is a streamlined version of the intuitions that (a) probabilistic dependencies always have a causal or nomological explanation and (b) that common causes and causal intermediaries screen off. Given (a), a causal dependency between x and y shows that x causes y , or y causes x or that x and y are effects of a common cause. It is arguable that the nomological connection that obtains between the paired particles in EPR phenomena provides a distinct explanation for probabilistic dependencies (Hausman [1998], pp. 247–52), but I shall ignore these complications here. If x and y are probabilistically dependent conditional on all the direct causes of x (and V is causally sufficient), then by (b) it cannot be the case that y causes x or that x and y are related only as effects of a common cause. So it must be the case that x

causes y . Although the causal Markov condition as formulated here is not true when there are causal cycles, related conditions can be defended (see Spirtes [1995]).

The faithfulness condition, on the other hand, expresses the assumption that probabilistic independencies result from causal structure, not from specific parameter values or the operation of latent causes. Birth-control pills impede pregnancy, which is one cause of blood clots. So, indirectly, birth-control pills impede the formation of blood clots. If birth-control pills were also a direct cause of blood clots of just the right strength (relative to these variables), it could turn out that blood clots were probabilistically independent of taking birth-control pills. Such a happenstance would violate the faithfulness condition, and as the example illustrates, such violations are not impossible. SGS argue that they will be rare and unstable. For instance, in this example, the probabilistic dependence between cause and effect—between birth control pills and blood clots—would reappear with a change in the fertility of those who do not take birth-control pills. It is not clear, however, how great a consolation this infrequency and instability are to practitioners, since apparent spurious independencies—that is, probabilistic dependencies that are too small to discriminate from zero—need not be so rare or unstable. David Freedman is highly critical of the faithfulness condition.

The simplest of SGS's algorithms finds equivalence classes of causal structures among a set of variables from information about their joint probability distribution. Although the following example works differently than the algorithm, it illustrates how causal conclusions can be drawn from probabilistic information, given the axioms and assumptions listed above. Suppose that one knows that the following probabilistic independencies obtain among the variables x , y , z , w , and r : (w, x) , (w, r) , (x, r) , (y, r) , and (w, z) are all unconditionally independent, and y and z are independent conditional on x . So the members of the first five pairs cannot be related as cause and effect or as effects of a common cause, and x must be either a common cause of y and z or a causal intermediary. If x is a causal intermediary, then either y causes x or z causes x . Suppose y causes x . Since y and w are not independent conditional on any subsets of V , they must be related as cause and effect. If w causes y , then, by transitivity, w causes x and, by faithfulness, w and x cannot be unconditionally independent. If y causes w , then x and w are effects of a common cause and so should be probabilistically dependent rather than independent. So y does not cause x . A similar argument shows that z does not cause x . So x is a common cause of y and z . Since x is independent of both w and r , faithfulness implies that x cannot be a cause—whether direct or indirect—of w or r . So it follows from causal sufficiency that w causes y and r causes z . The causal structure is:

$$w \rightarrow y \leftarrow x \rightarrow z \leftarrow r.$$

Using different methods, SGS's basic algorithm (the PC algorithm) draws

the same conclusion much more elegantly. It is based on the equivalence between conditional probabilistic dependence and a graph theoretic relationship explored by Judea Pearl and known as ‘d-separation’ (Pearl *et al.* [1990]). In particular, the PC algorithm would infer that x and w cause y and that x and r cause z from the surprising fact, that will be discussed later, that x and w are probabilistically dependent conditional on y and that x and r are probabilistically dependent conditional on z . Since the algorithm has been implemented in the TETRAD II program, this causal inference has been automated. When the causal sufficiency assumption fails (as it often does), the PC algorithm cannot be employed, but SGS have other algorithms that permit one to draw some conclusions concerning causal relations even when there are unmeasured common causes. I should emphasize (as SGS do) that their work borrows heavily from the work of other statisticians and computer scientists, especially from the work of Judea Pearl and his collaborators.

2 What is so provocative?

As even this thumbnail sketch makes clear, SGS seem to have accomplished something extraordinary. Rather than wallowing in the complexities of comparing the fit of alternative systems of regression equations or debating endlessly (and, in the absence of powerful social theories, inconclusively) what sort of causal relations ought theoretically to obtain, one merely tosses covariance information into the TETRAD II meat-grinder and out come neatly packaged causal conclusions. This seems too good to be true. Statisticians and philosophers are likely to object to different things, but many will share the reaction that SGS are playing some sort of conjuring trick. In different ways, Freedman, Humphreys, and Woodward all maintain that one needs to know more in order to reach causal conclusions from non-causal information about joint probability distributions.

Before turning to detailed criticisms, it is important to point out that the impression that SGS are conjuring causal conclusions out of almost thin air is misleading. The assumptions needed to apply the PC algorithm are compact and quickly stated: acyclicity, causal sufficiency, independent sources of variation, and the causal Markov and faithfulness conditions. But they are neither weak nor covert. Cyclic relations are common, at least at the level of the variables in many social theories. To know that a set of variables is causally sufficient is to have a great deal of causal knowledge—indeed much more than investigators typically have. Faithfulness may break down (or it may appear to break down). Although some of SGS’s other algorithms are less demanding, it should be evident that SGS are exploring the implications of strong causal assumptions rather than showing how one can make causal inferences without making causal assumptions.

3 Philosophical criticisms

I shall discuss three main philosophical questions concerning SGS's construction, which are raised by Paul Humphreys and James Woodward:

1. Does SGS's formal apparatus have a reasonable causal interpretation?
2. How should the 'causation' that SGS talk about be interpreted and how plausible is the causal Markov assumption?
3. Do the 'causal' relations SGS discover justify subjunctive conclusions concerning the results of interventions?

In his brief discussion, Humphreys discusses all three of these questions. The second and third, which are (as we shall see) closely related to one another, are also discussed by Woodward at much greater length. David Freedman's extensive critique, which I will not discuss here, reveals specific gaffs and oddities in some of SGS's case studies, and it presents cases in which causal sufficiency and faithfulness fail.

3.1 The causal interpretation of induced probabilistic dependencies

Humphreys points out that the PC algorithm relies on an implication of the causal Markov and faithfulness conditions that is hard to interpret causally. The causal Markov and faithfulness conditions imply that exogenous—that is, causally and probabilistically independent—causes of a common effect are probabilistically dependent conditional on their common effect. Humphreys gives Judea Pearl's example,

in which two coins are flipped and a bell rings if and only if both coins are the same. When we know that the bell rang (and when we know it did not) we can infer with certainty from the outcome of one coin toss what the other coin outcome was. This *induced dependency* between the coin outcomes is clearly epistemic—it is certainly not causal (pp. 254–5).

At first glance it may not be obvious what is bothering Humphreys. The fact that the coin flips are probabilistically dependent conditional on the bell's ringing merely licenses the inference that the coin flips are causes of the bell's ringing. As Humphreys is well aware, nobody is claiming that the coin flips bear any causal relationship to one another.

The difficulty, as I see it, is the following. The causal Markov condition implies that if there is an unconditional probabilistic dependency between x and y , then x causes y , y causes x , or that x and y are effects of a common cause. It says that if there is a probabilistic dependency between x and y conditional on all the direct causes of x , then x causes y . Probabilistic dependency is, as it were, the mark of causal relations. So the induced probabilistic dependency

between causes conditional on their joint effect ought to have a causal interpretation. Humphreys goes on to criticize SGS's unsuccessful effort to supply such an interpretation ([1993], pp. 72–3), and he argues that the only consistent way to understand this induced dependence is epistemic. Humphreys concludes that the supposedly causal relations that SGS infer from probabilistic relations are in fact only inferential relations (p. 256).

Although Humphreys is right to criticize SGS's discussion of such induced dependence, his conclusion is premature: the induced dependence can be given a causal interpretation. Since the mechanism that makes the bell ring when both coins land heads or tails is unspecified here, I prefer a more familiar example with the same structure. The upstairs hall light in my house (as in millions of others) is controlled by two switches, one upstairs and one downstairs. Assuming that the bulb has not burned out, the light is on if and only if both switches are up or down. The positions of the switches are unconditionally independent, but they are dependent conditional on the light being on (or off). Humphreys gives an epistemic reading to the phrase 'conditional on the light being on (or off)'. *Knowing* that the light is on or off, one can infer the position of the downstairs switch from knowing the position of the upstairs switch. The epistemic interpretation of the induced dependency is in this way built into the interpretation of conditioning. Similarly if one interprets conditioning on a common cause as *knowing* its state, then the 'induced independence' among its effects is merely epistemic.

What other reading of 'conditioning' is there? In conditioning on a variable, one is considering the state of affairs—possible world if you like—in which the variable's value is unchanging. In a possible world in which the common cause of x and y is fixed (and they each have their own independent sources of variation), there would be no probabilistic dependence between x and y . Not only would people be unable to make any inference about the value of x from knowing the value of y (given that they already know the value of their common cause), but there would *be* no causal connection between those values. Similarly, in a possible world in which the hall light is on and stays on, there would have to be some causal mechanism connecting the positions of the two switches so that as the downstairs switch is subjected to its own independent source of variation, the upstairs switch moves in tandem. As Paul Humphreys pointed out to me, the sources of variation in the position of each switch cannot be probabilistically independent of all other variables, and indeed the counterfactual relation between the position of the switches is not at the type-level asymmetrical. So the full articulation of this interpretation of induced probabilistic dependencies leads to complications that cannot be dealt with here. The crucial point is that Humphreys has not shown that probabilistic dependencies only indicate epistemic relations. They can sensibly be regarded as indicators of causal connections.

3.2 The interpretation of causation and the plausibility of the causal Markov condition

SGS disavow any attempt to analyse causation. Clark Glymour describes their research strategy as follows:

[T]here is no settled definition of causation, nor is it likely to be fruitful to insist on one before other questions are addressed. We know that causal relations are different from simple statistical associations and that the difference has to do with the fact that causal claims entail something subjunctive, something about what would happen under ideal interventions and manipulations. We have a wealth of cases from experimental design and from forms of causal explanation of nonexperimental data. And, most recently, we have an array of plausible proposals for axiomatizing the relations between causal and stochastic claims. Progress in understanding lies in exploring the mathematical consequences of various combinations of such axioms and in seeing what they imply for familiar cases and methods (p. 202).

This is perfectly sensible, and in my view SGS's achievements demonstrate the virtues of this procedure. But the procedure has its vices, too. In the sentences immediately before the quotation, Glymour suggests an analogy with the development of the theory of probability and statistical inference. Consider a different analogy, this time to social choice theory. One might interpret its many theorems as revealing the implications of the normative principles that social choice theorists were able to formalize, and no doubt some people read the literature this way. But it is at least as plausible to take the often surprising and remarkable results that have been proven as revealing the inadequacies of the formalizations of normative principles and the ambiguities within the normative principles themselves. In just the same way, one can question whether the causal Markov condition is equally plausible on every interpretation of causation; and one can take the formal results as casting doubt on the axioms rather than as showing how to make causal inferences.

One of the applications SGS develop in their 1993 book concerns the Armed Forces Qualification Test, which contains three subtests as components. SGS's TETRAD program correctly picks out these three subtests from a list of seven possibilities by examining the covariations among the scores. Freedman, Humphreys and Woodward all criticize SGS here. A simpler hypothetical example clarifies the difficulty. Suppose one has 1000 white boxes, numbered 1–1000, and 1000 red boxes also numbered 1–1000. These are randomly filled with different quantities of sand. Suppose one then has data reporting results of weighing each white box (w), weighing each red box (r) and weighing each pair of boxes with the same number (p). Since there are errors in all three measurements, each of the variables has its own source of variation. In

examining this data, one finds that w and r are unconditionally independent. Conditional on p , they are probabilistically dependent. Assuming causal sufficiency, PC will then conclude that w , and r cause p . This is indeed exactly what the causal Markov condition tells us directly, since p is dependent on w conditional on all the direct causes of w , and p is dependent on r conditional on all the direct causes of r . But rather than accepting the conclusion that r and w cause p , one ought to reject the claim that dependency conditional on direct causes is sufficient for causation. r and w are *parts* of p rather than causes of p . Humphreys (p. 259) and Woodward (p. 295f) make this criticism of SGS's treatment of the AFQT case and Freedman suggests the criticism by placing scare quotes around references to 'causes' and 'effects' (p. 133).

In this particular case, the cure is the same simple one that resolves related objections to counterfactual theories of causation. (The relation between the objection here and objections to counterfactual theories of causation is noted by Humphreys on p. 260.) Causation is a relation between distinct events. The possession by a pair of boxes of a certain mass at a specific place and time is not distinct from the possession there and then of a certain mass by the particular red box and the possession there and then of a certain mass by the particular white box. (The weighing events on the other hand may be distinct, and some of them are causally related as effects of common causes.) Provided that it is understood that the variables must be in the relevant sense distinct, there is no problem here for the causal Markov condition. Alternatively, the distinctness requirement can be built into the causal Markov condition. Since it may not always be obvious whether or not variables are distinct, one has another reason to regard the output of SGS's inference procedures as fallible.

Other problems may be more difficult to deal with. Here is one more example. As SGS as well aware, when one mixes data from two different populations, one may find a probabilistic dependency between x and y even though these variables are independent in the separate populations. This dependence will imply, misleadingly, that there is some sort of causal relationship between x and y . If one includes some variable z indicating from which population values of x and y are drawn, x and y will be independent conditional on z , and it will appear that z causes x and y or is a causal intermediary. Is this reasonable? Can one make sense of manipulating z as a means to influence x or y (Woodward, p. 294)?

3.3 Probabilistic dependency, causation, and manipulation

The most perplexing question raised by Humphreys and developed at some length by Woodward concerns the basis upon which one is supposed to be able to draw subjunctive conclusions (about what would happen if

one were to intervene) from nonexperimental data. In the example above in section 1, data on conditional and unconditional independencies plus various assumptions enabled us to conclude that the causal relation among w , y , x , z , and r is: $w \rightarrow y \leftarrow x \rightarrow z \leftarrow r$.

If one accepts the connection between causation and subjunctives concerning the results of (ideal) interventions, one should be justified in concluding that if one were to intervene and to set the value of x , the values of w and r would be unaffected, and the values of y and z would change in accordance with the quantitative relations relating the values of these variables to their causes. Of course one does not expect this to be the case for any possible value of x , since for some values of x the causal relation may break down. But the causal relations imply that there is some range of values of x for which x could in principle be used as a means to control the values of y and z .

How can facts about the probabilistic relations that obtain among the values of variables in a given population tell one about what would happen if there were an intervention that changed the value of x (Woodward pp. 186–93)? Recall, first, the assumptions needed to infer the causal relations: causal sufficiency (so that the relations cannot be due to a common cause), independent sources of variation (so that there will be variation of x within the population and probabilistic dependencies will not reflect *de facto* constancies), the causal Markov condition (needed to infer that x causes y and z), and the faithfulness condition (needed to infer that there are no other causal relations). Furthermore, the probability distribution is not, of course, given. It must be inferred from information concerning frequencies in samples, and to draw causal conclusions, one must take this inference to be sound. In addition, to move from claims about relations within a given distribution to subjunctive conclusions about what would be the case if one were to intervene and to set the value of x , one needs to assume that probabilistic dependencies among variables never arise by coincidence in the population as a whole, but always reflect the influence of causes. This assumption is of course already implicit in the causal Markov condition.

Can one then make the step? SGS have a ‘manipulation theorem’ that apparently says that accepting the causal Markov condition licenses this inference from probabilities to counterfactuals. Rather than accept this conclusion, one might, of course, take this result as a further argument against the causal Markov condition. (The theorem also depends on an additional controversial assumption that the manipulated and unmanipulated causal systems can be represented by a common ‘combined’ structure that includes an additional intervention variable that has an ‘on’ value when there is an intervention and an ‘off’ value when there isn’t one.) Woodward for one remains sceptical.

4 Conclusion

Whether SGS's algorithms and the computer programs that are based on them will lead to progress in social and epidemiological inquiries is difficult for me to say. I am not an empirical researcher, and so far there have been few attempts to apply their work to address empirical questions. Spirtes, Glymour, and Scheines have greatly clarified the principles upon which causal inferences have been made, and their work—like many of the other contributions to this volume—helps one to see how fragile and faulty those inferences often have been. Insofar as it helps people to understand what they have been doing, SGS's work can only do good. There is a risk, of course, of mindless applications; and it is an open question what will happen when practitioners of the statistical arts put SGS's techniques to work on the variables and data sets they possess. Only time will tell whether SGS's techniques will in fact be more informative and reliable than seat-of-the-pants unsystematic theorizing.

Judgement of the philosophical value of their work does not have to wait. There is in my mind no doubt that SGS have made a major philosophical contribution. The formalization of claims concerning causation and the mathematical investigation of their consequences is (of course) no substitute for philosophical thought concerning the nature of causation and its relation to explanation, counterfactuals, interventions, processes, and time. But fortunately the mathematical inquiries do not need to wait for a resolution to the philosophical difficulties. By revealing the implications and ambiguities in philosophical theses concerning causation, such inquiries can help philosophers to make progress with the many recalcitrant metaphysical difficulties to which the notion of causation gives rise. By its actual contribution, the work of SGS decisively establishes this possibility. *Causality in Crisis?* is a wonderful introduction to this exciting area of philosophy.

*London School of Economics and
University of Wisconsin-Madison*

Acknowledgements

I would like to thank Clark Glymour, Paul Humphreys, David Papineau, Richard Scheines, and James Woodward for comments on an earlier draft. None of them fully agrees with the views expressed in this review.

References

- Hausman, Daniel [1998]: *Causal Asymmetries*, New York: Cambridge University Press.
- Pearl, Judea, Geiger, Dan, and Verma, Thomas [1990]: 'The Logic of Influence Diagrams', in R. M. Oliver and J. Q. Smith (eds), *Influence Diagrams, Belief Nets and Decision Analysis*, New York: John Wiley & Sons.

Spirtes, Peter [1995]: 'Directed Cyclic Graphical Representations of Feedback Models', in Philippe Besnard and Steven Hanks (*eds*), *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence*, San Mateo, CA: Morgan Kaufmann Publishers.

Spirtes, Peter, Glymour, Clark, and Scheines, Richard [1993]: *Causation, Prediction, and Search*, New York: Springer-Verlag.