4

Modularity: It Can—and Generally Does—Fail*

NANCY CARTWRIGHT

London School of Economics
University of California, San Diego

1 Introduction

This paper pursues themes developed in my recent book *The Dappled World: A Study of the Boundaries of Science.* (Cartwright 1999). The book is a Scotist book—in accord with the viewpoint of Duns Scotus. It extols the particular over the universal, the diverse over the homogeneous and the local over the global. Its central thesis is that in the world that science studies, differences turn out to matter. Correlatively, universal methods and universal theories should be viewed with suspicion. We should look very carefully at their empirical justification before we adopt them.

The topic in this volume is causality; I shall defend a particularist view of our subject. Causal systems differ. What is characteristic of one is not characteristic of all and the methods that work for finding out about one need not work for finding out about another. I shall argue this here for one specific characteristic: *modularity.* Very roughly, a system of causal laws is modular in the sense I shall discuss when each effect in the system has one cause all its own, a cause that causes it but does not cause any other effect in the system. On the face of it this may seem a very special, probably rare,

* Research for this paper was supported by a grant from the Latsis Foundation.

Stochastic Causality.
Maria Carla Galavotti, Patrick Suppes and Domenico Costantini (eds.).

situation. But a number of authors currently writing on causality suppose just the opposite. Modularity, they say, is a universal characteristic of causal systems. I shall argue that they are mistaken.

2 What Is Modularity?

Behind the idea that each effect in a causal system[1] should have a cause of its own is another idea, the idea that each effect in the system must be able to take any value in its range consistent with all other effects in the system taking any values in their ranges. There are two standard ways in which people seem to think this can happen; it will be apparent that different senses of *able* are involved.

In the first place, a second collection of causal systems very similar to the first may be possible in which all the laws are exactly the same except for the laws for the particular effect in question. In the new systems these laws are replaced by new laws that dictate that the effect take some specific value, where the systems in the collection cover all the values in the range of that effect.

This interpretation clearly requires that we be able to make sense of the claim that an alternative set of laws is possible. For my own part I have no trouble with this concept: in *The Dappled World* I argue that laws are not fundamental but instead arise as the result of the successful operation of a stable arrangement of features with stable capacities. Nevertheless, I do not see any grounds for the assumption that the right kind of alternative arrangements must be possible to give rise to just the right sets of laws to make modularity true. At any rate this way of securing modularity is not my topic in this paper.

The second way in which modularity might obtain is when each effect in the system has a cause all of its own that can contribute to whatever its other causes are doing to make the effect take any value in its range. This is the one I will discuss here. I will also in this paper restrict my attention to systems of causal laws that are both linear and deterministic. In this case the commitment to modularity of the second kind becomes a commitment to what I call "epistemic convenience".

An *epistemically convenient linear deterministic system* is a system of causal laws of the following form[2]

---

[1] I shall use "causal system" to refer to a set of causal laws and "causal structure" to refer to a set of hypotheses about causal laws.

[2] Somewhat more accurately, I should say "a system of laws generated by laws of the following form", for I take it that causal laws are transitive. For a more precise formulation, see Cartwright (2000).

$$x_1 \text{ c=} u_1$$
$$x_2 \text{ c=} a_{21}x_1 + u_2$$
$$\vdots$$
$$x_n \text{ c=} \Sigma a_{nj}x_j + u_n$$

plus a probability measure $P(u_1,\ldots,u_n)$, where

i. there are no cross restraints among the u's[3] and the u's are probabilistically independent of each other;

ii. for all j, $\text{Prob}(u_j = 0) \neq 1$.

The symbol "c=" shows that the law is a causal law. It implies both that the relation obtained by replacing "c=" with "=" holds and that all the quantities referred to on the right-hand side are causes of the one on the left-hand side.

These systems are epistemically convenient because they make it easy to employ randomized treatment/control experiments to settle questions about causality using the method of concomitant variation. I will explain in more detail below, but the basic idea can be seen by considering the most straightforward version of the method of concomitant variation: to test if x_j causes x_e and with what strength, use u_j to vary x_j while holding fixed all the other u's, and look to see how x_e varies in train. Conditions i) and ii) guarantee that this can be done.

A number of authors from different fields maintain that modularity[4] is a universal characteristic of causal systems. This includes economic methodologist Kevin Hoover[5], possibly Herbert Simon (1977), economists T. F. Cooley and Stephan LeRoy (1985), Judea Pearl (2000) in his new study of counterfactuals, James Woodward (1997), Daniel Hausman (1998) and Daniel Hausman and James Woodward (1999) jointly in a paper on the causal Markov condition. I aim to show that, contrary to their claims, we can have causality without modularity. I focus on the second kind of modularity here in part because that is the kind I have found most explicitly

---

[3] See n. 16.

[4] Or some closely related doctrine. Much of what I say can be reformulated to bear on various different versions of a modularity-like condition.

[5] See his defense of the invariance of the conditional probability of the effect on the cause in Hoover (forthcoming). In this discussion Hoover seems to suppose that there always is some way for a single cause to vary and to do so without any change in the overall set of laws. At other places, I think, he does not assume this. But he does speak with approval of Herbert Simon's way of characterizing causal order, and Simon's charcterization requires the possibility of separate variation of each factor.

defended. Hence I shall be arguing that *not all causal systems are epistemically convenient.*[6]

3 The Method of Concomitant Variation

We *say* that the method of concomitant variation is a good way to test a causal claim. But can we show it? For an epistemically convenient system we can, given certain natural assumptions about causal systems. That is one of the best things about an epistemically convenient system—we can use the method of concomitant variation to find out about it.

I shall not give the proof here, but rather describe some results we can show.

Here are the assumptions I shall make about linear deterministic systems of causal laws:

A_1: *Functional dependence.* Any causal equation re presents a true functional relation.

A_2: *Anti-symmetry and irreflexivity.* If q causes r, r does not cause q and q does not cause q.

A_3: *Uniqueness of coefficients.* No effect has more than one expansion in the same set of causes.

A_4: *Numerical transitivity.* Causally correct equations remain causally correct if we substitute for any effect any function in its causes that is among nature's causal laws.

A_5: *Consistency.* Any two causally correct equations for the same effect can be brought into the same form by substituting for effects in them functions of their causes given in nature's causal laws.

A_6: *Generalized Reichenbach principle.* No quantities are functionally related unless the relation follows from nature's causal laws.

The result I shall describe says very roughly that when the underlying linear deterministic system is an epistemically convenient one, then a causal hypothesis is correct iff the method of concomitant variation says it is so. In order to express this more precisely, we shall have to know what the form of the causal hypotheses in question are, what it is for hypotheses of this

[6] The authors mentioned here all have slightly different views, formulated and defended differently and with different caveats. I apologize for lumping them all together. Clearly not all the remarks I make are relevant to every view. In fact I will focus on a very specific form of the claim for universal modularity. Nevertheless, most of what I say can be translated to apply to other forms of the claim.

form to be causally correct and what it is to pass the test for concomitant variation.

The usual hypotheses on offer when we suppose the underlying causal system to be linear and deterministic[7] are in the form of regression equations:

R: $x_k \; c= \Sigma a_{kj}x_j + \Psi_k$, for $\Psi_k \perp x_j$ for all j,
where $x \perp y$ means that $<xy> = <x><y>$.

What exactly does this hypothesis claim? I take it that the usual understanding is this: every quantity represented by a named variable (an "x") on the right-hand side is a genuine cause of the quantity represented on the left-hand side, and the coefficients are "right". The random variable "Ψ" represents a sum of not-yet-known causes that turn R into a direct representation of one of the laws of the system. So I propose to define correctness thus: an equation of the form R: $x_k \; c= \Sigma a_{kj}x_j + \Psi_k$ $(1 \leq j \leq m)$, for $\Psi_k \perp x_{j,}$ is *correct* iff there exist $\{b_j\}$ (possibly $b_j = 0$), $\{q_j'\}$ such that $q_k \; c= \Sigma a_{kj}q_j + \Sigma b_j q_j' + u_k (1 \leq j \leq m)$, where q_j does not cause q_j'. (This last restriction ensures that all the omitted factors are causally antecedent to or "simultaneous" with those mentioned in the regression formula. Note, x_j represents q_j.)

Now let us consider *concomitant variation*. In an epistemically convenient linear deterministic system, the value of the x's are fixed by the u's, and the u's can vary independently of each other. The core idea of the method is to take the concomitant variation between x_c and x_e when u_c varies while all the other u's are fixed as a measure of the coefficient of x_c in nature's equation for x_e.

To state the relevant theorems we shall need some notation. Let $\Delta j(\alpha)x_n =df\, x_n(u_1 = U_1, \ldots , u_{j-1} =U_{j-1}, u_j = U_j+\alpha, u_{j+1} = U_{j+1}, \ldots , u_m = U_m) - x_n(u_1 = U_1, \ldots , u_{j-1} =U_{j-1}, u_j = U_j, u_{j+1} = U_{j+1}, \ldots , u_m= U_m)$. Then we can prove[8]

Theorem 1. A (true) regression equation $x_k \; c= \Sigma_{j=1}^{k-1} a_{kj}x_j + \Psi_k$, is causally correct iff for all values of α and J, $1 \leq J \leq k$, $\Delta_J(\alpha)x_k = \Sigma a_{kj}\Delta_J(\alpha)x_j$; i.e. iff the equation predicts rightly the differences in x_k generated from variations in any right-hand-side variable.

Notice, however, that this theorem is not very helpful to us in making causal inferences because it will be hard to tell whether an equation has indeed predicted the differences rightly. That is because we will not know

[7] Or when we are prepared to model the system as linear and deterministic for some reason or another.

[8] For a proof of the three theorems see Cartwright (2000, Ch. 3, also forthcoming in *Philosophy of Science*.). The formalization and proofs are inspired by the work of James Woodward on invariance, which argues more informally for a more loosely stated claim. I am aiming to make these claims more precise in these theorems.

what $\Delta_J(\alpha)x_j$ should be unless we know how variations in u_J affect x_j and to know that we will have to know the causal relations between x_J and x_j. So in order to judge whether each of the x_j affects x_k in the way hypothesized, we will have already to know how they affect each other. If we happen to know that none of them affect the others at all, we will be in a better situation, since the following can be trivially derived from the previous theorem:

Theorem 2. A (true) regression equation for x_k in which no right-hand side variable causes any other is causally correct iff for all α and J, $\Delta_J(\alpha)x_k = a_{kJ}\Delta_J(\alpha)u_J$.

We can also do somewhat better if we have a complete set of hypotheses about the right-hand-side variables. To explain this, let me define a *complete causal structure* that represents an epistemically convenient linear deterministic system with probability measure, as a triple, $<X = \{x_1,\ldots,x_n\}, \mu, CLH>$, where μ is a probability measure over X and where the causal law hypotheses, CLH, have the following form:

$$x_1 \ c= \Psi_1$$
$$x_2 \ c= a_{21}x_1 + \Psi_2$$
$$\vdots$$
$$x_n \ c= \Sigma_{j=1}^{n-1} a_{nj}x_j + \Psi_n,$$

with $\Psi_j \perp x_k$, for all $k < j$. In general $n < m$, where m is the number of effects in the causal system. Now I can formulate

Theorem 3. If for all x_k in a complete causal structure, the $\Delta_J(\alpha)x_k$ that actually obtains equals $\Delta_J(\alpha)x_k$ *as predicted by the causal structure* for all α and J, $1 \leq J \leq n$, then all the hypotheses of the structure are correct.

I take it that it is the kind of facts recorded in these theorems that make epistemically convenient systems so desirable, so that we might wish—if we could have it—for all causal systems to be epistemically convenient. But is it sensible to think they are? In the next section I will give some obvious starting reasons for thinking the answer must be "no".

4 Three Peculiarities of Epistemic Convenience

To notice how odd the requirement of epistemic convenience is, let us look first at some ordinary object whose operation would naturally be modeled at most points by a system of deterministic laws—for instance a well-made toaster like the one in Figure 1[9]. The expansion of the sensor due to the heat

[9] Figure 1 is drawn by Emily Cartwright following an explanation and illustration of the functioning of a toaster in Macaulay (1988).

The Toaster

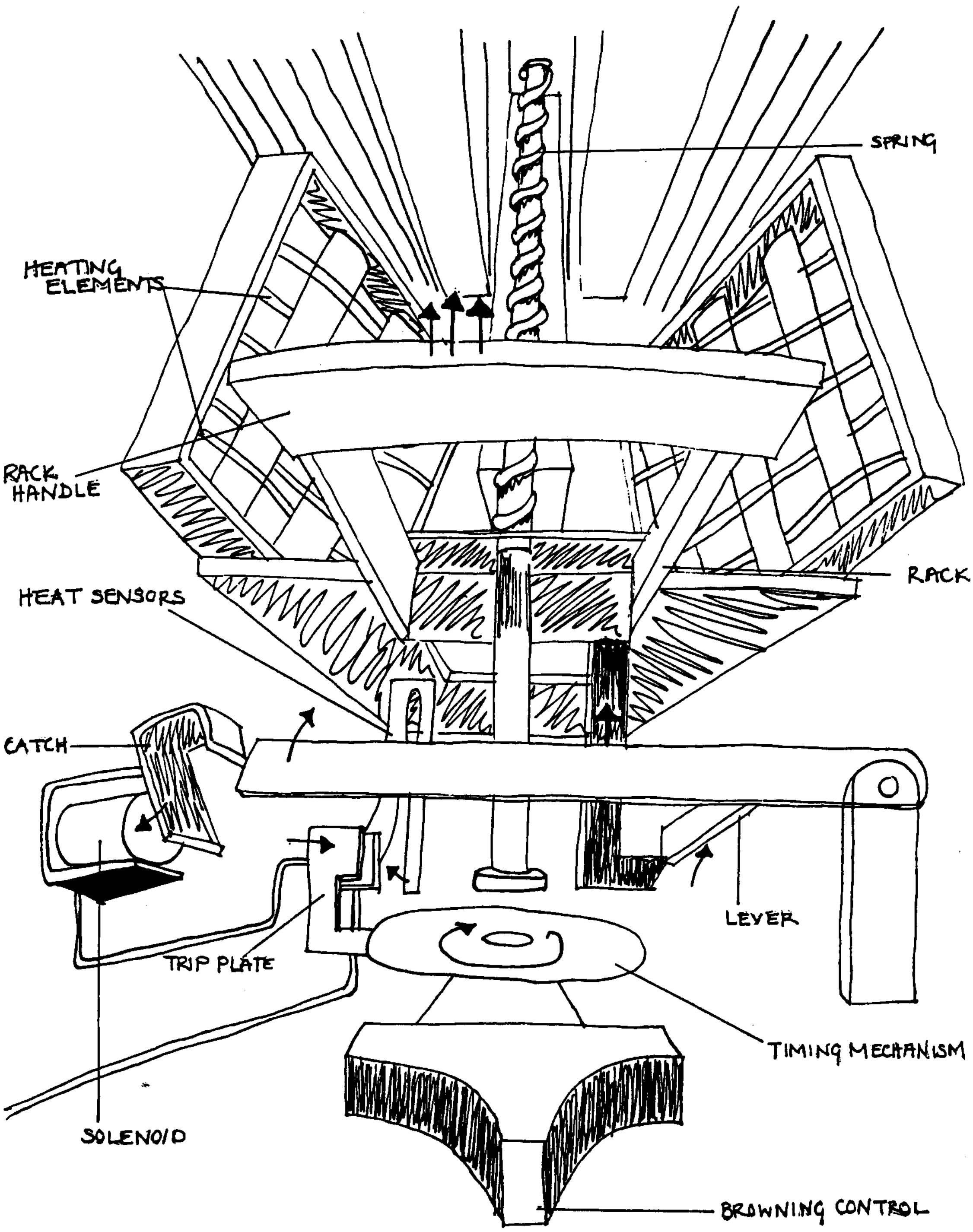

Figure 1.

produces a contact between the trip plate and the sensor. This completes the circuit, allowing the solenoid to attract the catch, which releases the lever. The lever moves forward and pushes the toast rack open.

I would say that the movement of the lever causes the movement of the rack. It also causes a break in the circuit. Where then is the special cause that affects only the movement of the rack? Indeed, where is there space for it? The rack is bolted to the lever. The rack must move exactly as the lever dictates. So long as the toaster stays intact and operates as it is supposed to, the movement of the rack must be fixed by the movement of the lever to which it is bolted.

Perhaps, though, we should take the bolting of the lever to the rack as an additional cause of the movement of the rack? In my opinion we should not. To do so is to mix up causes that produce effects within the properly operating toaster with the facts responsible for the toaster operating in the way it does; that is, to confuse the causal laws at work with the reason those are the causal laws at work.[10] But even if we did add the bolting together at this point as a cause, I do not see how it could satisfy conditions i) and ii). It does after all happen as part of the execution of the overall design of the toaster, and hence it is highly correlated with all the other similar causes that we should add if we add this one, such as the locating of the trip plate and the locating of the sensor.

The second thing that is odd about the demand for modularity is where it locates the causal nexus. It is usual to suppose that the fact that C causes E depends on some relations between C and E.[11] Modularity makes it depend on the relation between the causes of C and C: C cannot cause anything unless it itself is brought about in a very special way.

Indeed, I think that Daniel Hausman embraces this view: "...people ... believe that causes make their effects happen and not vice versa. This belief is an exaggerated metaphysical pun, which derives from the fact that people can make things happen by their causes. This belief presupposes the possibility of intervention and the claim that not all the causes of a given event are nomically connected to one another" (Hausman 1998, 272).

This is a very strong view that should be contrasted with the weaker view, closer (on my reading) to that of Hume, that the concept of causation *arises* because "people can make things happen by their causes", but that this condition does not constitute a truth condition for causation. The

---

[10] For a more complete discussion of this point, see the distinction between nomological machines, on the one hand, and the laws that such machines give rise to, on the other, in Cartwright (1999).

[11] Or perhaps, since *C* and *E* here pick out types and not particular events, "between C-type events and E-type events".

weaker view requires at most that sometimes a cause of a cause of a given effect vary independently of all the "other" causes of that effect; it does not take epistemic convenience as universal.[12] In my opinion the weaker view only makes sense as an empirical or historical claim about how we do in fact form our concepts, and about that, we still do not have a reliable account. The stronger view just seems odd.

Thirdly, the doctrine seems to imply that it is impossible to build a bomb that cannot be defused. Nor can we make a deterministic device of this sort: the correct functioning of the mechanisms requires that they operate in a vacuum; so we seal the whole device in a vacuum in such a way that we cannot penetrate the cover to affect one cause in the chain without affecting all of them. Maybe we cannot build a device of this sort—but why not? It does not seem like the claim that we cannot build a perpetual motion machine. On the doctrine of universal epistemic convenience we either have to say that these devices are indeed impossible, or that what is going on from one step to the next inside the cover is not causation, no matter how much it looks like other known cases of causation or passes other tests for causation (such as the transfer of energy/momentum test or the demand that a cause increase the probability of its effects holding fixed a full set of other causes).

Given that the claim to empirical convenience as a universal condition on causality has these odd features, what might motivate one to adopt it? Three motivations are ready to hand: we might be moved by operationalist intuitions, or by pragmatist intuitions or we might be very optimistic about how nicely the world is arranged for us. I will take up each in turn.

5 Motivations for Epistemic Convenience: 'Excessive' Operationalism

This is a hypothesis of Arthur Fine's[13]: Advocates of modularity conflate the truth conditions for a causal claim with conditions which *were they to*

[12] It is also surprising that Hausman focuses on the supposition of (something like) epistemic convenience as a necessary condition, but does not stress the equally problematic matter of the possibility of choice. We all know the classic debate about free will and determinism: it looks as if people cannot make things happen by their causes unless the causes of the causes are not themselves determined by factors outside our will, and that in turn looks to preclude universal determinism. If that should follow, it would not trouble me, but many advocates of modularity also defend the causal Markov condition—which I attack—on the grounds of universal determinism. Moreover, the need for *us* to cause some of the causes at least some of the time seems equally necessary whether one takes the strong view that Hausman maintains or the weaker view—which does not require epistemic convenience as a truth condition—that I described as closer to Hume.

[13] Conversation, May, 2000, Athens, Ohio.

obtain would make for a ready test. As we have seen, a central feature of deterministic systems that are epistemically convenient is that we can use the simplest version of the method of concomitant variation within them: to test "x_c cause x_e", consider situations in which x_c varies *without variation in any 'other' causes of* x_e and look for variation in x_e. I think this is particularly plausible as a motivation for economists. Economists in general tend to be very loyal to empiricism, even to the point of adopting operationalism. For instance, they do not like to admit preferences as a psychological category but prefer to use only preferences that are revealed in actions.

In general, versions of operationalism that elevate a good test to a truth condition are in disfavour. Still we need not dispute the matter here, for, even were we disposed to this kind of operationalism in the special case at hand, it would not do the job. Simple concomitant variation is no better test than many others—including more complicated methods of concomitant variation. So operationalism will not lead us to limit causal concepts to systems that admit tests by simple concomitant variation at the cost of other kinds of systems. In particular, the simple method does not demand any 'less'[14] background knowledge than tests using more complicated versions of concomitant variation, which can be performed on other kinds of deterministic systems, nor knowledge of a different kind.

Let me illustrate. We will continue to look at linear deterministic systems and we will still assume that all exogenous factors are mutually unconstrained: there are no functional relations among them.[15] And we will still test for causal relations by the method of concomitant variation.

Imagine then that we wish to learn the overall strength, if any, of x_1's capacity to affect x_e, where we assume we know some cause u_1 that has a known effect (of, say, size b_1) on x_1 and whose variation we can observe. In the general case where we do not presuppose epistemic convenience, every candidate for u_1 may well affect x_e by other intermediaries, say $x_2,\ldots,x_n$, as well. Suppose the overall strength of its capacity to affect x_j is b_j and of x_j to affect x_e is c_j.

We aim to compare two different situations, which are identified by the values assigned to the u's: $S = \langle U_1,U_2,\ldots,U_m\rangle$ and $S' = \langle U_1' \ U_1,U_2,\ldots,U_m\rangle$, where the u's constitute a complete set of mutually unconstrained exogenous factors that determine x_e. Then

[14] I put "less" in scare quotes because I do not mean to get us involved in any formal notions of more and less information.

[15] If they are constrained, re-express all of them as appropriate functions of a further set of mutually unconstrained factors. Notice that this has no implications one way or another about whether, for example, two endogenous factors share all the same exogenous causes. The point of this is to allow that the exogenous factors *can* vary independently of each other.

$$x_e\text{'}-x_e = \Sigma_{j=1}{}^{m} b_j c_j (U_1\text{'}-U_1)$$

or, letting $(\Sigma_{j=2}{}^{m} b_j c_j)/b_1$ =df A,

$$(x_e\text{'}-x_e)/b_1(U_1\text{'}-U_1) = c_1+A.$$

Now here is the argument we might be tempted to give in favour of epistemically convenient systems. If we have an epistemically convenient system, A = 0, so $c_1 = x_e\text{'}-x_e/b_1(U_1\text{'}-U_1)$. Otherwise we need to know the value of A. So we need less antecedent knowledge if our systems are epistemically convenient.

But clearly the last two sentences are mistaken: A = 0 is just as much a value of A as any other; to apply the method of concomitant variation, we need to know (or be willing to bet on) the value of A in any case. *Sometimes* there may be some factor u_1 for which it is fairly easy to know that its effect on x_e by routes other than x_1 is zero. This for example, is, in my opinion, the case with J.L. Mackies's famous hypothesis that the sounding of the end-of-workday hooters in Manchester brings the workers out onto the streets in London. Here we know various ways to make the hooters in Manchester sound of which we can be fairly confident that they could not get the workers in London out except via making the Manchester hooters hoot.[16]

But equally, sometimes we may know for some exogenous factors that do affect x_e by routes other than x_1 what the overall strength of that effect is—if, for instance, we have data on variations in x_e given variations in u_1 when the route from u_1 to x_1 is blocked.

Let us review some of the prominent facts we would need to know for a brute force application of the method of concomitant variation, as I have described it, in a linear deterministic system. To test "x_1 causes x_e with strength c" we need to know

1. of a factor u_1 that it is exogenous to the system under study, that it causes x_1 and with what strength it does so;
2. of a set of factors that they are exogenous, that they are mutually unconstrained, and that together, possibly including u_1, they are sufficient to fix x_e but not sufficient to fix u_1;
3. what would happen to x_e in two different situations for which the values of the exogenous factors described in 2. do not vary, except for the value of u_1, which does vary;
4. the overall strength of u_1's capacity to affect x_e by other routes than by causing x_1.

---

[16] For a discussion, see Cartwright (1989).

My point here is that we need to know (or find a way around knowing) all of this information whether or not the system is epistemically convenient.

Why then have I called these special kinds of systems "epistemically convenient" for use of the method of concomitant variation if we need to know (or find our way around knowing) "the same amount" of information to use the method whether the system is epistemically convenient or not? Because when the system is epistemically convenient, it is a lot easier to use randomized treatment/control experiments. That is why I have called these systems "epistemically convenient"; and it is one of the chief arguments James Woodward (2000, p. 10) gives in favour of the claim that causal systems should be epistemically convenient:

> A manipulationist approach to causation explains the role of experimentation in causal inference in a very simple and straightforward way: Experimentation is relevant to establishing causal claims because those claims consist in, or have immediate implications concerning, claims about what would happen to effects under appropriate manipulations of their putative causes. In other words, the connection between causation and experimentation is built into the very content of causal claims.

Randomized treatment/control experiments provide us with a powerful tool to find our way around knowing large chunks of information we otherwise would need to know. For the point at issue in this paper, we need to be clear about which features of the stock experimental structure help with which aspects of our ignorance.

Randomization allows us to finesse our lack of knowledge of the kinds of facts described in 2. above. When we are considering the effect of u_1 on x_e, we generally do not know a set of "other" exogenous factors sufficient to fix x_e. But a successful randomization ensures that they will be equally distributed in both the treatment and the control groups. Hence there will be no background correlations between these other factors that might confound our results. Observing the outcome in the two groups allows us to find out (roughly[17]) the information we look for in 3.: what happens under variation in u_1.

But notice that randomization and observation do these jobs whether or not the system is epistemically convenient. Epistemic convenience matters

[17] The experiment does not allow us to tell what happens for any two specific situations (i.e., any specific choice of values for the u_j) but only certain coarser facts. For instance, if u_1 is causally unanimous across *all* situations (i.e., it is either causally positive across all, or causally negative or causally neutral), for a two-valued outcome, x_e, it can be shown that the probability of x_e in the treatment group is respectively greater than, less than or equal to that in the control group iff u_1 is causally positive, negative or neutral with respect to x_e.

because we were trying to find out, not about the effects of u_1, but rather about the effects of x_1. In the case I described above, where epistemic convenience fails, u_1 has *multiple capacities*: it can affect x_e differently by different routes. We are interested only in its effect via x_1, which we shall use to calculate the effect of x_1 itself.[18] Randomization does not help with this problem. Just as in the brute force application of concomitant variation, we need either to find a cause of x_1 which we know has no other way of affecting x_e, or we need to know the overall effect via other routes in order to subtract it away.

The placebo effect is a well-known example of this problem. Getting the experimental subjects to take the medicine not only causes them to have the medicine in their bodies. It can also affect recovery by producing various psychological effects—feeling cared for, optimism about recovery, etc.

This is a good example to reflect on with respect to the general question of how widespread epistemically convenient systems are. How do we canonically deal with the placebo effect? We give the patients in the control group some treatment that is outwardly as similar to the treatment under test as possible but that is known to have no effect on the outcome under study.

That is, we do not hunt for yet another way to get the medicine into the subjects, a way that does not affect recovery by any other route. Rather we accept that our methods of so doing may affect recovery in the way suggested (or by still other routes) and introduce another factor into the control group that we hope will just balance whatever these effects (if any) may be. Ironically then, the standard procedure in these medical experiments does not support the claim that there is always a way to manipulate the cause we want to test without in any other way affecting the outcome. Epistemic convenience definitely makes randomized treatment/control experiments easier, but there are vast numbers of cases in which we do not rely on it to hold.

6 Motivations for Epistemic Convenience: 'Excessive' Pragmatism

This is a hypothesis raised by the students in my Ph.D. seminar on causality in economics at LSE: Advocates of modularity elevate a plausible answer to the question "Of what use to us is a concept of causation?" into a truth condition. This motivation is explicitly acknowledged by Daniel Hausman

---

[18] It may be just worth reminding ourselves, so as not to confuse the two issues, that x_1 itself may have multiple capacities with respect to x_e. Simple randomized treatment / control experiments do not disentangle these various capacities but rather teach us about the overall effect of x_1 on x_e.

(1998, 96-97) in defense of a similar condition to the one we are investigating:

> What do people need causal explanations or a notion of causation for? Why isn't it enough to know the lawlike relations among types? Because human beings are actors, not just spectators. Knowledge of laws will guide the expectation of spectators, but it does not tell actors what will result from their interventions. The possibility of abstract intervention is essential to causation...

My remarks here are identical to those about operationalism. Whether or not we wish to adopt the pragmatic justification as a truth condition, it will not do the job of defending modularity as a truth condition. Consider the same example as above. The conditions for using variations in u_1 to produce variations in x_1 and thereby to obtain predictable variations in x_e are much the same as the conditions for testing via concomitant variation.

To know what we will bring about in x_e by manipulating u_1 it is not enough to know just the influence of u_1 on x_1 and of x_1 on x_e. We also need to know the overall influence of u_1 on x_e by all other routes. Knowing that the size of influence by other routes is zero is just a special case. Whatever its value, if we know what the value is we can couple this knowledge with our knowledge of the influence of u_1 via x_1 to make reliable predictions of the consequences of our actions. So we do not need modularity to make use of our causal knowledge.

There is, however, a venerable argument for a different conclusion lurking here. If we are to use our causal knowledge of the link from x_1 to bring about values we want for x_e, it seems that some cause or other of x_1 must not itself be deterministically fixed by factors independent of our wishes: there must be some causes of x_1 that we can genuinely manipulate. But again, whether or not this is a good argument, it does not bear on modularity. To make use of our knowledge of the causal link between x_1 and x_e we may need a cause of x_1 that we can manipulate; but that does not show that we need a cause we can manipulate without in any other way affecting x_e.

7 Motivations for Epistemic Convenience: 'Excessive' Optimism

This is my hypothesis. Life becomes easier in a number of ways if the systems we study are economically convenient. Statistical inference of the strengths of coefficients in linear equations can become easier in well-known ways. So too can causal inference, in ways I have discussed here. And, as we shall see, Judea Pearl can provide a very nice semantics for counterfactuals as well as for a number of distinct causal notions. Wishful thinking can lead us to believe that all systems we encounter will meet the conditions that make life easier. But wishful thinking must be avoided here, or we will be led into the use of methods that cannot deliver what we rely on them for.

I think we can conclude from these considerations that these three motivations do not provide strong enough reason to accept universal economic convenience. What positive arguments then are on offer on its behalf?

8 For and Against Economic Convenience

8.1 Hausman's Defense

Daniel Hausman points out that the cause we focus on is not generally the complete cause. A complete cause will include both helping factors and the absence of disturbances. Even if effects share the causes we normally focus on (e.g., in the toaster in Figure 1, the breaking of the circuit and the moving of the rack are both caused by the motion of the lever), they will not share all of these other factors, Hausman maintains.

Disturbing factors. This claim seems particularly plausible with respect to disturbing factors. Most of the effects we are modeling here are fairly well separated in time and space. So it seems reasonable to expect that some things that might disturb the one would not disturb the other. This seems promising for the thesis of, if not universal, at least widespread, epistemic convenience. But there is a trouble with disturbing factors: often what they do is to disrupt the relation between the causes and the effect altogether. To salvage economic convenience, they need instead to cooperate with the causes adding or subtracting any spare influence necessary to ensure that the effect can take all the values in its allowed range. So they do not seem to satisfy reliably the conditions for epistemic convenience.

Helping factors. Return to the toaster. The motion of the lever causes the motion of the rack. That of course depends on the fact that the lever is bolted solidly to the rack: if the lever were not bolted to the rack, the lever could not move the rack. Could we not then take the fact that the lever is bolted to the rack to be just what we need for the special cause of the motion of the rack, a cause that the motion of the rack has all to itself?

I think not, for a number reasons:

1. As I urged in section 4, the fact that the two are bolted together is not one of the causes within the system of causal laws but rather part of the identification of what that systems of laws applies to, and this identification matters. We do not, after all, seek to know what the causal law is

that links the movement of levers *in general* with the movement of racks of the right shape to contain toast. Surely there is no such law. Rather we want to know the causal relation, if any, between the movement of the lever and the movement of the rack *in a toaster of this particular design*. Without a specific design under consideration, the question of the causal connection, or lack of it, between levers and racks is meaningless.

2. Let us, however, for the sake of argument, admit as a helping cause in the laws determining the motion of the rack the fact that the lever and rack are bolted together. My second worry about calling on helping factors like this to save epistemic convenience depends on the probability relations these factors must bear to each other. In section 4, I queried whether these factors would be probabilistically independent of each other. Here I want to ask a prior question. Where is the probability distribution over these factors supposed to come from and what does it mean?

   We could consider as our reference class toasters meeting the specific set of design requirements under consideration. Then the probabilities for all of these "helping factors" being just as they are could be defined and would be 1. Independence would be trivially obtained, but at the cost of the kind of variation we need in the values of the u's to guarantee, via our theorems, that concomitant variation will give the right verdicts about causality.

   Alternatively the reference class could be the toasters produced in a given factory following the designated design. Presumably then there would be some faults some time in affixing the lever to the rack so that not all the u's would have probability 1. But will the faults be independent? If not this reference class, then what? It will not do to have a make-believe class, for how are we to answer the question: if the attachment of the lever to the rack *were* to vary, what would happen to the rack? We need some other information to tell us that—most usually, I would suppose, knowledge of the causal connections in the toaster! And if not that information exactly, I bet it would nevertheless be information sufficient to settle the causal issue directly, without detour through concomitant variations.

3. The third worry is about the range of variation. For the theorems to guarantee the reliability of the method of concomitant variation, we need u's that will take the cause under test through its full range relative to the full range of values for the other causes. Otherwise there could be blips—the causal equation we infer is not true across all the

values but depends on the specific arrangement of values we consider. Will the factors we pick out have a reasonable range of variation? This remark applies equally well to disturbing factors.

4. Last I should like to point out two peculiarities in the way people often talk about the factors designated by the u's. Often they are supposed not only to represent the special causes peculiar to each separate effect but also all the "unknown" factors we have not included in our model. But if they are unknown, they can hardly be of use to us as handles for applying the method of concomitant variation. And if epistemically convenient systems are not going to be of epistemic convenience after all, why should we want them? I realize that the issue here is not supposed to be whether we want systems of this kind, but rather whether we have them. Still in cases like this where the answer is hard to make out, the strategy should be to ask "What depends on the answer". That is the reasonable way to establish clear criteria for whether a proffered answer is acceptable or not.

   The second peculiarity arises from talking of the u's as a "switch" that turns the cause to different values. Often it is proposed that the switch is usually "off" yet could be turned on to allow us to intervene. This raises worries about the independence requirements on the u's again. Why should that kind of factor have a probability distribution at all, let alone one that renders it independent of all the other switch variables?

8.2 Pearl's Defense

Judea Pearl supposes that modularity holds in the semantics he provides for singular counterfactuals. He claims that, without modularity, counterfactuals would be ambiguous.[19] So modularity must obtain wherever counterfactuals make sense. This will double as an argument for universal modularity if we think that counterfactuals make sense in every causal system.

Pearl assumes modularity of the first kind, where alternative causal systems of just the right kind are always possible, but I can explain something of how the semantics works using the epistemically convenient systems we have been studying here. We ask, for instance, in a situation where $x_j = X_j$ and $x_k = X_k$, "Were $x_j = X_j+\Delta$, would $x_k = X_k+a_{kj}\Delta$?" The question may be thought ambiguous because we do not know what is to stay fixed as x_k varies. Not so if we adopt the analogue of Pearl's semantics for our epistmically convenient system. In that case u_j must vary in order to produce the required variation in x_j and all the others u's must stay the same.

19 14 March, 2000, seminar presentation, Dept. of Philosophy, University of California at San Diego.

The semantics Pearl offers is very nice, but I do not see how it functions as an argument that counterfactuals need modularity. The counterfactuals become unambiguous just because Pearl provides a semantics for them and because that semantics always provides a yes-no outcome. Any semantics that does this will equally make them unambiguous.

Perhaps we could argue on Pearl's behalf that his is the *right* semantics, and it is a semantics that is not available in systems that are not epistemically convenient. Against that we have all the standard arguments that counterfactuals are used in different ways, and Pearl's semantics—like others—only accounts for some of our uses. We should also point out that we do use, and seem to understand, counterfactuals in situations where it is in no way apparent that the causal laws at work are epistemically convenient.

I think one defense Pearl may have in mind concerns the connection between counterfactuals and causality. Consider a very simple case where one common cause, v, is totally responsible both for $x_1$and x_2 and no u_1 is available to vary x_1 independently of v.[20]

It is easy to construct a semantics, similar indeed to the one Pearl does construct, that answers unambiguously what would happen to x_2 if x_1 were different. This semantics would dictate that we vary v to achieve the variation in x_1. Then of course x_2 would vary as well. So it would be true that were x_1 to be different, x_2 would be. And that seems a perfectly reasonable claim for some purposes. But not of course if we wish to read singular causal facts from our counterfactuals. So Pearl could argue that his semantics for counterfactuals connects singular counterfactuals and singular causal claims in the right way. And his semantics needs modularity. So modularity is a universal feature wherever singular causal claims make sense.

Laying aside tangled questions about the relations between singular causal claims and causal laws, which latter are the topic of this paper, I still do not think this argument will work.

We could admit that for an epistemically convenient system Pearl's semantics for counterfactuals plus the counterfactual-causal links he lays out will give correct judgments about causal claims. We could in addition admit that causal claims cannot be judged by this method if the system is not epistemically convenient. All this shows is that methods that are really good for making judgments in one kind of system need not work in another kind.

More strongly, we could perhaps somehow become convinced that no formal semantics for causal claims that works, as Pearl's does, by trans-

[20] For Pearl this would mean that there was no alternative causal system possible that substituted the law "Let $x_1 = X_1$" for the law connecting x_1 and v.

forming a test into a truth condition, will succeed across all systems of laws that are not epistemically convenient. That would not show that there are no causal laws in those systems, but merely that facts about causal laws are not reducible to facts about the outcomes of tests.

9 Conclusion

Modularity is not a universal feature of deterministic causal systems, nice as it would be were it universal. Part of my argument for this conclusion depends on asking of various factors, such as the fact that the toaster rack is bolted to the lever, "Are these really causes?" I argued that they are not because they cannot do for us what we want these particular kinds of causes to do. In this case what we want is a guarantee that if we use these factors in applying the method of concomitant variation, the results will be reliable.

I think this is the right way to answer the question. We should not sit and dispute whether a certain factor in a given situation is really a cause, or what causality really is. Rather we should look to whether the factor will serve the purposes for which we need a "cause" on this occasion. That means, however, that for different purposes the very same factor functioning in the very same way in the very same context will sometimes be a cause and sometimes not.

That is all to the good. Causality is a loose cluster concept. We can say causes bring about their effects, but there are a thousand and one different roles one factor can play in bringing about another. Some may be fairly standard from case to case; others, peculiar to specific structures in specific situations. Causal judgements, and the methods for making them reliably, depend on the use to which the judgment will be put. I would not, of course, want to deny that there may be some ranges of cases and some ranges of circumstances where a single off-the-shelf concept of causality, or a single off-the-shelf method, will suffice. But even then, before we invest heavily in any consequences of our judgments, we need strong reassurance both that this claim is true for the ranges supposed and that our case sits squarely within those ranges.

That of course makes life far more difficult than a once-and-for-all judgment, a multipurpose tool that can be carried around from case to case, a tool that needs little knowledge of the local scene or the local needs to apply. But it would be foolhardy to suppose that the easy tool or the cheap tool or the tool we happen to have at hand must be the reliable tool.

References

Cartwright, N. (1989) *Nature's Capacities and Their Measurement.* Oxford: Oxford University Press.

Cartwright, N. (1999) *The Dappled World: A Study of the Boundaries of Science.* Cambridge: Cambridge University Press.

Cartwright, N. (2000) *Measuring Causes: Invariance, Modularity and the Causal Markov Condition.* London: CPNSS Measurement in Physics and Economics Discussion Paper Series, LSE.

Cooley T. F. and LeRoy, S. F. (1985) A theoretical macroeconometrics. *Journal of Monetary Economics,* 16, 283-308.

Hausman, D. (1998) *Causal Asymmetries.* Cambridge: Cambridge University Press.

Hausman, D. and Woodward, J. (1999) Independence, invariance and the causal Markov condition. *British Journal for the Philosophy of Science,* 50, 1-63.

Hoover (forthcoming) *Causality in Macroeconomics.* Cambridge: Cambridge University Press.

Macaulay, D. (1988) *The Way Things Work.* Boston: Houghton Mifflin.

Pearl, J. (2000) *Causality.* Cambridge: Cambridge University Press.

Simon, H. (1977) Causal ordering and identifiability. In H. Simon (Ed.) *Models of Discovery and Other Topics in the Methods of Science.* Boston: Reidel, 53-105.

Woodward, J. (1997) Explanation, invariance and intervention. *British Journal for the Philosophy of Science* 64, 26-41.

Woodward, J. (2000) Causation and manipulation. ms, California Institute of Technology.