The Statistical Nature of Causation

David Papineau

1 Introduction

Causation is a macroscopic phenomenon. The temporal asymmetry displayed by causation has no counterpart in the fundamental dynamics of the microscopic world. So causation must somehow emerge from the underlying dynamics, along with other macroscopic phenomena like entropy increase and the arrow of radiation.

In this paper I shall offer an account of causation that accounts for this emergence. I shall show how macroscopic events fall into lawlike patterns that are subject to the random influences of microscopic processes. It is the probabilistic independence of these random influences that constitutes these laws as temporally asymmetric and therefore causal.

I shall approach these issues by considering the “causal inference” techniques that allow causal relations to be inferred from sets of observed correlations. I shall show that these techniques are best explained by a reduction of causation to structures of equations with probabilistically independent exogenous terms. This probabilistic independence imposes a recursive order on these equations and a consequent distinction between dependent effect variables and independent cause variables. This then accounts for the temporal asymmetry of causation, since in reality the independent variables in any equation always occur before their dependent variables.[[1]](#footnote-1)

Philosophers in the tradition of David Lewis have sought in various way to account for the asymmetry of causation in terms of the “asymmetry of overdetermination”. These accounts are consonant with the analysis that I shall offer, but from my perspective they start the story far too late. The asymmetry of overdetermination is not something to be assumed, but itself a macroscopic phenomenon that itself demands explanation. My analysis will have the virtue of displaying the asymmetry of overdetermination as a consequence of the asymmetric nature of causation.[[2]](#footnote-2)

2 Explaining Causal Inference

For over a hundred years non-experimental scientists have been inferring asymmetrical causal conclusions from correlational data. Curiously, none of the main philosophical theories of causation cast any light on why these techniques work. What in the nature of causation allows such inferences to proceed? As far as I know, nobody working on counterfactual, or regularity, or process, or dispositional theories of causation so much as asks this question.

The only philosophers who have addressed this issue have been those in the minority tradition of “probabilistic theories” of causation that attempt to reduce causal relationships directly to correlational ones. In the middle of the last century Hans Reichenbach (1956), I.J. Good (1961-2) and Patrick Suppes (1970) all offered variations on this theme, and more recently Wolfgang Spohn (2001), Clark Glymour (2004), and Gerhard Schurtz and Alexander Gebharter (2016) and also myself (2001) have drawn on the analysis of “Bayesian networks” to develop more sophisticated versions of this strategy. However, as I shall show, this tradition ties causation too closely to correlations. Because of this, it cannot cope with “faithfulness failures” where correlations are misleading about causal links, nor is it able explain the relationship between probabilistic causal connections and single-case actual causation.

A less ambitious version of this probabilistic analysis is associated with the so-called “interventionist” approach to causation (Woodward 2003, 2016). In effect, this approach disavows any reductive ambitions and instead contents itself with articulating some non-reductive connections between causes and correlations. While this unsurprisingly allows the interventionist approach to offer some explanation of the causal inference techniques (Hausman and Woodward 1999), it still faces difficulties with faithfulness failures and with single-case causation.

My strategy in this paper will be to offer a different reductive analysis. The idea goes back to H. A. Simon and others in the 1950s and 60s (Simon 1953, Blalock 1964). It seeks to reduce causation to underlying structural equations with probabilistic independent error terms rather than directly to surface correlations. While this idea is often enough aired by practising non-experimental scientists, it has been largely ignored by philosophers (though see Cartwright 1989). I shall show that this approach holds the key both to the success of the correlational inferential techniques and to the way that causes are arranged asymmetrically in time.

3 Bridge Principles

Let me quickly illustrate the kind of causal inference techniques at issue. Suppose that good examination results (E) are positively correlated[[3]](#footnote-3) with attendance at private rather than state schools (S), but this correlation disappears when we control for parental income (P)--as it is said, P *screens off* the correlation, in the sense that E is no longer more likely given S once we hold P fixed. We conclude that, despite the prima facie indication that private schools causally influences examination results, this correlation is in fact “spurious” in the sense that schools S don’t affect examination results E after all; rather both are effects of the common cause parental income, as in the following causal *directed acyclic structure* (“DAS” henceforth)[[4]](#footnote-4):

(1) P

S

E

Inferences like these are natural enough, but they do not wear their rationale on their sleeve. Recent work in the tradition of “Bayesian networks” has regimented the principles behind them. (See, for example, Spirtes et al 1993, Pearl 2000, Peters et al 2017.) In this section I shall lay out these principles. After that I shall consider what grounds them.

It will be convenient for what follows to say that two variables X and Y are *causally linked* if X causes Y (possibly indirectly via intermediaries), or Y causes X (again possibly indirectly), or X and Y have a (possibly indirect) common cause—but not if X and Y only have a common effect.

One issue that arises at this point is what is meant by saying that one generic variable “causes” another, in the sense in which, say, parental income might cause examination results. Presumably such claims are some kind of generalization over particular cases in which definite values of these variables bear some relation to each other. But how to fill this out is by no means obvious. Still, it will be convenient for our purposes to take the generic notion of cause as read for the time being. By the end of the paper I will be able to explain it.

Inferring causes from correlations hinges on two kinds of principles—I shall call them “Bridge Principles” henceforth. On the one hand are a pair of principles licensing moves *from* correlations *to* causes, and on the other a pair licensing moves *from absence* of correlation *to absence* of causes. (The former pair are often presented together as the “Causal Markov Condition” and the latter pair as the “Faithfulness Condition”. But it will be more illuminating to unpack them as follows.)

First and simplest is the *Linkage Principle*:

(2) If two variables are correlated, then they must be causally linked.

And to this can be added a *Conditional Linkage Principle*:

(3) If two correlated variables remain conditionally correlated after we control for other variables {X}, then they must be casually linked by one or more paths that do not go via {X}.

These two principles specify that correlations always indicate a causal link: correlated variables must either be cause and effect or they must have a common cause. Moreover, correlations that persist even after controlling for other variables indicate casual links that that by-pass those controlling variables.[[5]](#footnote-5)

Now for the converse principles.

The *Unlinkage Principle* says:

(4) If two variables are uncorrelated, then they are not causally linked.

And a further *Conditional Unlinkage Principle* says:

(5) If two correlated variables are screened off by other variables {X}, then they are not causally linked by any chains of variables that do not contain any of {X}.

These two principles now tell us that variables that aren’t correlated can’t be causally linked: they can’t cause each other or have a common cause. Moreover, two variables that cease to be correlated when we control for other variables can’t be linked in any ways that by-pass those controlling variables.[[6]](#footnote-6)

Together these principles suffice to fix a causal order among variables displaying a sufficiently rich set of correlations.

Of course they don’t fix a causal order among any set of correlated variables. To revert to our original schools (S)-examinations (E)-parental income (P) example, the pairwise correlations between all three variables plus the screening-off of the S-E correlation by P is in itself consistent, not only with the possibility suggested earlier, weher P causes both S and E, as in (1), but also with

(6) P

S

E

and

(7) P

S

E

Still, this indeterminacy would be resolved if we were given some further variable A that was only correlated with S and another B that was only correlated with E. Then the only causal structure consistent with the bridge principles would be

(8)

P

A

B

S

E

This illustrates a result that can be proved in full generality. Whenever the correlations between some set of variables do not allow the bridge principles to fix their causal relationships uniquely, there will always be possible correlations involving further possible variables that will so suffice (Theorem 4.6 Spirtes et al 1993 p 94).

Let me observe at this point that the bridge principles fix causal order, when they do, without resorting to any information about the temporal ordering of the relevant variables. Yet we can expect that, when they do so fix causal order, the variables identified as causes will always precede their effects in time. If this is so, this must be because this temporal ordering of variables is implicit in the correlations displayed by causally related sets of variables. The arrangement of correlations is itself asymmetrically distributed in time. This augurs well for the project of understanding how the temporal asymmetry of causation can emerge in a world with a symmetric fundamental dynamics.

4 Causal Sufficiency

The points made so far might suggest the reductive idea that there is nothing more to causal relations than the patterns of correlation from which they can be inferred via the bridge principles. In effect, this would be to view the bridge principles as necessary truths that encode the way that causal relations are implicit in correlational structures.

I shall explore this idea further in the next section. But first an immediate issue must be addressed. I have shown that the bridge principles can determine a causal order given the correlations *among a given set of variables*. But this leaves it open that a causal order so determined might be overturned if the set of variables were expanded. This possibility clearly threatens the idea that causal relations are nothing over and above the correlational patterns that imply them. After all, our aim is to analyse the nature of *real* causal relations, not of *apparent* causal relations *relative to* an arbitrary selection of variables.

This worry is by no means an idle one. Imagine that the correlations among {A, B, S, P, E} determine

(8)

P

A

B

S

E

in the way described in the last section. Despite this unique determination of a causal order by the correlations, it remains perfectly possible that in reality P does not cause S, but rather both are effects of some further G. (Imagine, not entirely implausibly, that parental income P per se has no effect on school type S, but that both are effects of grandparental income G.) In that case, the correlations among {A, B, S, P, E} would remain just as observed, but the bridge-principle driven conclusion that P causes S would simply be wrong, and the true structure would be

(9) G

P

A

B

S

E

However, while this is a real danger, it is significantly limited. To see why, note that bringing in extra variables will not itself alter the correlations among the variables in some original set. For example, expanding the analysis by including G won’t stop P being systematically correlated with S. At most, extra variables will screen off correlations that weren’t screened-off in the original variable set. If we now apply the bridge principles to the expanded set of variables, any such newly exposed screening off will then indicate that causal links presented as direct by the original set are in fact only indirect causal links.

Now, indirect causal links are of two kinds—either one variable causes another via an intermediary, or two variables have a common cause. In the first kind of case, expanding our variable set will not really have overturned any causal conclusion, since it will only have shown that some causal link proceeds via intermediaries, as would in any case have been assumed. So it is only the second kind of case, where the extra variable turns out to be a common cause of two variables in the original set, that the casual verdicts delivered by the original correlations will be reversed.

This now shows that there will be no overturning of verdicts delivered by the bridge principles as long as we have a *causally sufficient* set of variables, in the sense of a set that does not omit any variables that turn out on expansion to be common causes of variables included in the original set. And this now opens the way to the reductive project once more. A revised reductive suggestion would now be that causal relations are nothing over and above those patterns of correlation that imply them, courtesy of the bridge principles, by in any causally sufficient set of variables. (Would not the need to specify *causal* sufficiency here render this suggestion inadmissibly circular as a reduction of causation? But this specification can be finessed away. We can simply say causal relations are nothing over and above the patterns of correlation that imply them in sets of variables whose verdicts are not overturned by the inclusion of further variables.)

5 Probabilistic Theories of Causation

The original probabilistic theories of causation, proposed in the middle of the twentieth century by Reichenbach (1956), Good (1961-2) and Suppes (1970), were all versions on this theme:

(10) An earlier X causes of a later Y if and only if they are positively correlated and this correlation is not screened off by any yet earlier Z.

But there are obvious drawbacks to this formulation. For a start, it appeals to the temporal order in its analysis of causal order and thus abandons the search for an independent explanation of why causal relations are asymmetric in time. Moreover, this formulation is ill-suited to accommodate various complex causal structures, as when there are two common causes of two correlated effect variables, with neither cause therefore fully screening the correlation among the effects.

The regimentation of the bridge principles in the Bayesian network tradition allows probabilistic theories of causation to by-pass these two difficulties. As we have seen, the bridge principles are capable of determining a causal order among any complex set of correlated variables, and they can do so without assuming any information about the temporal ordering of those variables. This then opens the way for theories which take the nature of causation to be implicitly captured by the bridge principles. On this account, there is nothing more to nothing more to causal relations than the patterns of correlation from which they can be inferred via the bridge principles. (See, for example, Spohn 2001, Papineau 2001, Glymour 2004, and Schurtz and Gebharter 2016.)

Can this implicit reduction be transformed into an explicit analysis of causation? The writers just cited do not attempt this, but the theory offered by Daniel Hausman in his *Causal Asymmetries* (1998) can be adapted for this purpose. Hausman himself does not offer an explicit definition of causation in terms of correlations, because of the “failures of faithfulness” that I shall address in the next section, but if we put that issue to one side for the moment, we can adapt his analysis and say:

(11) X causes Y if and only if X is correlated with Y and everything correlated with X is correlated with Y and something correlated with Y is not correlated with X.

The basic idea behind this reduction is that the effects in correlated cause-effect pairs are distinguished from the causes by having probabilistically independent sources of variation, while correlated joint effects of common causes are distinguished from cause-effect pairs by *both* having independent sources of variation.

If we make the assumption that effects do always have such independent sources of variation, then the reductive claim (11) follows from the bridge principles (2)-(5). The need to add this assumption of independent sources of variation to the bridge principles is a reflection of the point, made in the section before last, that not every set of correlations suffices for the bridge principles to determine a causal order, but for any set of correlations there is a possible expanded set that will suffice for this.

6 Failures of Faithfulness

Unfortunately, we cannot rest with this kind of probabilistic reduction of causation. It focuses on the surface symptoms of causation, rather than its underlying nature.

One way to see this is to note that the causal claims treated by this reduction will leave us without answers to various questions of single-case causation. Suppose, as is no doubt the case, that the bridge principles in conjunction with actual empirical correlations establish that *smoking causes lung cancer*. Now suppose that Joe Bloggs smokes and gets lung cancer. Did his smoking cause his lung cancer? It depends. Even if it is true that “smoking causes lung cancer” in the sense inferred from the bridge principles and the correlations, Joe’s genetic make-up might prevent cigarettes from harming him and he might have acquired his lung cancer from asbestos exposure instead.

This shows that there must be more structure to causal relations than is captured by the kind of causal claims at issue in the probabilistic reduction. I shall propose an account of this extra structure in the next section.

Associated with this point are doubts about the metaphysical status of the bridge principles themselves. The probabilistic reduction argues that they are necessities whose truth falls out of the nature of causation itself. But this is doubtful.

All the bridge principles have been questioned in the literature. As it happens, I think that the objections to the Causal Markov Condition are misplaced, and that the truth of this does indeed fall out of the underlying nature of causation, as I shall show in section 9 below. But the Faithfulness Condition has a far less secure status, and even if generally true is by no means metaphysically guaranteed.

Let me first briefly indicate why I view objections to the Causal Markov Condition as misplaced. Both of its corollaries, the Linkage and Conditional Linkage Principle, have been queried in the literature.

The standard counter-examples to the Linkage Principle are time-series correlations which point to no causal linkage, like the matched values of London bread prices and Venice water levels over the years. As a range of writers have pointed out, however, the Linkage Principle can be qualified so as to exclude correlations with this type of non-standard time-series construction (Sober 2001, Hoover 2003, Zhang and Spirtes 2014).

Then there are putative counter-examples to the Conditional Linkage Principle. These are cases of correlations which supposedly fail to be screened off when causal intermediaries are controlled for. Wesley Salmon’s pool balls (1984) and Nancy Cartwright’s polluting factory (2002) are well-known examples. Here the standard response is that the cases omit a full specification of the intermediaries and that that the correlations would disappear if this were rectified. (Let me put to one side until section 12 possible counterexamples to the Conditional Linkage Principle involving quantum measurements on separated entangled particles. These raise different issues.)

Now for the Faithfulness Condition. The problem cases here involve a cause influencing an effect via two separate pathways, one producing the effect and the other preventing it, with the two influences cancelling out, and so leaving us with a zero correlation between the cause and the effect. The classical example, due to Hesslow (1976), involves birth control pills B exerting a direct positive influence on the formation of thromboses T, but also working to prevent T by preventing pregnancies P which themselves operate as a positive cause of thromboses. If the two influences cancelled out precisely, then B and T would end up quite uncorrelated with thromboses just as common among those taking birth control pills as those who don’t. (This would be a counterexample to the simple Linkage Principle. But similar examples to Conditional Linkage can easily be constructed—for instance, just imagine that some C is a common cause of both B and T.)

(12) B

-

+ P

+

T

Case like these would be clear violations of the Faithfulness Condition. B and T are uncorrelated, even though causally linked. If we applied the bridge principles to this case, we would conclude that the uncorrelated B and T are independent causes of pregnancies P, since they are both correlated with it—which by hypothesis is not the real causal structure.

Now such “failures of faithfulness” are clearly highly unlikely, given the way they depend on an exact cancelling out of influences. Still, this is little consolation to those defending a neo-probabilistic theory of causation. For that theory requires the Faithfulness Condition to be a necessary consequence of the nature of causation. And cases like Hesslow’s, however unlikely they might be, just don’t seem metaphysically impossible. It certainly doesn’t look as if there is anything in the nature of causation to stop two causal paths cancelling out exactly and leaving us with a null association between cause and effect. I conclude that we need to probe deeper to understand the connection between the bridge principle and the underlying nature of causation. (Before moving on, it is worth observing that, while exact failures of faithfulness might be highly unlikely, and so only of abstract philosophical significance, *approximate* failures of faithfulness are common enough in the real world, and so a practical concern to those non-experimental scientists who need to start their investigations by using finite sample data to estimate correlations.)

7 Structural Equations

To get to the bottom of the connection between causation and correlation, we need to switch to a different approach, namely the “structural equations” tradition that played a central role in econometrics and sociometrics in the middle of the last century. To revert to our earlier example, this tradition would no doubt have dealt with the covariation of schools (S), examination results (E) and parental income (P) by positing a set of linear equations like this[[7]](#footnote-7):

(13.1) P = eP

(13.2) S = aP + eS

(13.3) E = bP + cS + eE

In such a set of equations, the e-terms that appear only on the right-hand sides are called “exogenous variables” and represent further influences beyond those explicitly included in the study. The other variables on the right-hand side of each equation are called its “independent variables”, and the variable on the left-hand side is its “dependent variable”.

A set of such equations is *recursive* if it can be ordered so that no term appears on the right-hand side as an independent variable unless it has appeared in a previous equation as a dependent variable. In such a case, the equation set will have a by a directed acyclic structure. In our example, this would be

eP

P

eS

S eE

E

The regression coefficients a, b, c attaching to the independent variables measure the extent to which the dependent variables vary in response to changes in those independent variables. They capture how much, if at all, the dependent variable “wiggles” when a given independent variable “wiggles” and the other independent variables are held constant.

In our example, we supposed that examination results E don’t co-vary at all with schooling S once parental income P is held constant. So then the regression coefficient c will be zero, and the equations will have the simpler structure:

(14.1) P = eP

(14.2) S = aP + eS

(14.3) E = bP + eE

eP

P

eS

S eE

E

I have now started representing sets of equations like (13) and (14) by directed acyclic structures--DASs. But note that this is a different kind of DAS from those considered earlier. Where the earlier DASs represented *causal* structures, these new DASs simply represent the order in which variables appear in structures of equations like (13) and (14).

Still, it is both standard and natural to interpret these new equational structures causally, and assume that one variable is a cause of another just in case it is an ancestor of it in the structure of equations—that is, to take our new equation-DASs also to be causal-DASs.

Still, it is not clear that anything said so far *justifies* such N interpretation. After all, if the equations are just equations, what is to stop us rearranging terms so that the dependent variables become independent and vice versa? For example, what is to stop us rewriting the equations (14) as follows?

(15.1) S = e\*S

(15.2) P = 1/aS + e\*P

(15.3) E = bP + eE

(with e\*S = aeP + eS , e\*P = -eS/a).

This would then give us this directed acyclic equation structure:

e\*P

P

e\*S

S e\*E

E

And, if we were to interpret this structure causally, it would now present S as a cause of P, and P as a cause of E, and S as having no direct causal influence on E except via P.

8 Independent Exogenous Terms

Even so, an approach to causation in terms of structural equations has the resources to meet this challenge. The key idea is that the ordering of variables in a set of equations will capture causal structure just in case *the exogenous variables are probabilistically independent*. This requirement promises to decide between the alternative causal hypotheses offered by equations (14) and (15). If P causes both S and E, then the exogenous terms in the former but not latter equations will be independent, whereas if S causes P which causes E, then the reverse will be true.

This idea was commonplace among econometricians and sociometricians in the middle of the last century. As they saw it, the requirement of exogenous independence wasn’t some arbitrary add-on to systems of structural equations, but built into their use as a tool for prediction and explanation. To see why, just consider the equations

(14.1) P = eP

(14.2) S = aP + eS

which presents school type S as a function of parental income P. Given some value for P, we would naturally use this to infer that expected school funding is aP. But note how this inference hinges on the implicit premise that the extra variation in S is independent of what value of P we have. That is why we can estimate S on the basis of knowing P even while being quite ignorant of the value of es.

To drive the point home, observe that this kind of inference doesn’t work the other way around. Consider, instead of (14.1-14.2), the rearranged

(15.1) S = e\*S

(15.2) P = 1/aS + e\*P

If es was independent of P in the original (14.1-14.2), then e\*P won’t be independent of S in (15.1-15.2) – remember that e\*P = -eS/a – and so now we can’t infer that 1/aS will be the average value of P given some value of S. The way P varies around 1/aS will be different for different values of S, and will depend on how P itself is distributed. (For example, if the median of P is below the average, as we would expect for parental income, then the expected value of P for positive S will generally be less than 1/aS.)

In my view, the probabilistic independence of exogenous terms in recursive systems of structural equations holds the key to causal direction. This kind of independence means that the values of the dependent variables are due to influences that can be factorised into independent sources. Since the exogenous terms are probabilistically independent of each other, since they display no systematic tendency to vary in concert, they are constituted as causally unlinked. The dependent variables, by contrast, are built up from these factorizable influences, and are thereby constituted as their effects.

So far I have illustrated the idea in a maximally simple case with just two non-exogenous variables. But the idea that independent exogenous terms constitute causal structure can be applied in more complex cases. To illustrate, let us imagine, contrary to our supposition so far, that schools S do after all exert an extra influence on examination results E, in addition to any direct influence from parental income P. The relevant equations would then be the earlier (13):

(13.1) P = eP

(13.2) S = aP + eS

(13.3) E = bP + cS + eE

eP

P

eS

S eE

E

If the sequence of exogenous terms is independent, then the values of P are fixed by one set of factors eP, the values of S are fixed by P plus another probabilistically independent set of factors eS, which thus constitutes S as an effect of P and eS--and finally the values of E are fixed by the values of P and S, which are themselves now correlated, and by yet another set of factors eE which are probabilistically independent of both P and S. This last independence thus constitutes E as an effect of all of P, S and eE.

The underlying idea, then, is that every dependent variable will have an associated exogenous variable that is independent of the other variables it is dependent on. This reflects the assumption, presupposed by the explicit probabilistic reduction of causation offered earlier in section 5, that effects will always have sources of variation that are independent of their other causes. But now we have built this assumption into a more structured framework that will prove better suited to deal with failures of faithfulness and single-case causation.

I have been using linear regression analysis to illustrate the idea that causal structure might depend on independent exogenous terms. But the idea can happily be generalised to other structures of deterministic equations. We needn’t restrict ourselves to linear equations, nor to real-valued variables.

Suppose we have any set of variables X1, . . . Xn and exogenous terms, E1, . . . En, possibly with values that might be dichotomous, or determinable, as well as quantitative in some way; and suppose we have a set of recursive deterministic equations over these variables of the form

(16) Xi = F(X1, . . . Xi-1, Ei).

Then in general we can view these equations as capturing causal structure on the assumption that the exogenous terms are probabilistically independent.

Just as with the linear regression examples I have used, the independence of the error terms is naturally viewed as imbuing the whole system of equations with causal structure. Each dependent variable X has its values fixed by its independent variables and its own exogenous term. The latter adds some X-specific variation to the value for X fixed by the independent variables, and so constitutes X as an effect of the terms on the right-hand side of its equation.

This now suggests the following reductive analysis of causation:

(17) X causes Y if an only if it is an ancestor of Y in a recursive structure of deterministic equations with independent exogenous terms.

On this account, a structure of causes and effects is nothing over and above a structure of variables in a system of deterministic equations with independent exogenous terms. In effect, this suggested analysis of causation combines a *regularity* theory of causal *covariation* with a *statistical* account of causal *direction*. We start with a set of deterministic equations. These specify how certain variables covary deterministically in a lawlike way.[[8]](#footnote-8) But this covariation is itself undirected. The covariation specified by the equations would remain the same if we reordered the equations to switch which sides the variables appeared on. The causal direction is then added to the covariation by the requirement that the exogenous terms be probabilistically independent of each other.

The reduction I have now offered presupposes that all the variables within any causal structure will be connected by deterministic equations. Quantum mechanics gives us reason to doubt that this is true. In section 12 below I shall offer a modified version of the reduction that will accommodate quantum indeterminism within causal structures. But before that it will be illuminating to explore the implications of the deterministic version of the reduction.

9 The Bridge Principles Recovered

The reduction of causation proposed above casts a new light on the bridge principles that underlie inferences from correlations to causation. Instead of seeing the correlations as providing the substance of causation, as on probabilistic theories of causation, we can now view them as offering indirect evidence for the way variables feature in systems of structural equations. On this account, when we use the bridge principles to infer that X causes Y from correlations, we are in fact inferring that X is an ancestor of Y in a system of structural equations.

Crucial in this connection is a mathematical theorem that I shall call the call the “*Independence-Markov Result*”. Suppose as before that we have a system of dependent variables X1, . . . Xn, with probabilistically independent exogenous variables, E1, . . . En, and recursive deterministic equations over these variables of the form Xi = F(X1, . . . Xi-1, Ei). Then the system will display this Markov property:

(18) Any variable will be probabilistically independent of every other variable (apart from its descendants) conditional on its parents (where “parent” and “descendant” signify the obvious relations in the DAS of the relevant equations). (Pearl 2000 Theorem 1.4.1.)

This result is obvious enough. Any two dependent variables that owe their values to disjoint sets of exogenous variables will inherit their independence from the independence of those exogenous variables. Putting it the other way round, two dependent variables will be correlated only if in the equations one descends from the other or they have a common ancestor. Moreover, if a variable does so descend from another, or shares a common ancestor with it, then any correlation between them will disappear if we hold fixed its parents, because any residual variation in the two variables will then again derive from disjoint sets of independent exogenous variables.

Note that as it stands this Independence-Markov Result says nothingabout *causes* as such. It is a straightforward mathematical claim about the joint probability distribution imposed on all the variables in a system of deterministic equations by the requirement that the exogenous terms be independent. Still, when we combine this theorem with the proposed reduction of causation to equation systems with exogenous independence, then it does imply the *Causal* Markov Condition—any *causal* structure will satisfy the Markov property—an therewith our earlier Linkage Principle—correlated variables must be causally linked—and Conditional Linkage Principle—if correlated variables remain correlated when we control for some further variable, then they must be linked by a route that does not involve that further variable.

So the analysis of causation I am recommending can fully account for the use of the Linkage and Conditional Linkage Principles to infer causal conclusions from correlational premises. It is noteworthy, though, that it does not simply *posit* that causal structures will satisfy these conditions, as did the probabilistic theories of causation considered earlier. Rather it *derives* this from the proposed deeper analysis of causation in terms of deterministic equations with independent exogenous variables.

It is also noteworthy that the Faithfulness Condition does *not* follow from the proposed reduction. That is just as it should be. As we saw earlier, it is highly implausible to suppose that the Faithfulness Condition is built into the metaphysical nature of causation. True, we can generally expect variables that are linked in a system of equations with exogenous independence to be correlated. The equational links plus the background independence will generally lead to the variables co-varying. In certain special cases, however, a specific cancelling-out of coefficients will mean that equationally linked variables display overall no correlation.

Recall Hesslow’s example in which birth control pills affect thromboses both directly but also via preventing pregnancies which themselves conduce to thromboses. Within the structural equations framework, this set-up might be realised by the following equations.

(19.1) B = eB

(13.2) P = aB + eP

(13.3) T = bB + cP + eT

Now in this case, and indeed in all cases with this equational structure, there will be no correlation between B and T if the coefficients cancel out exactly and a + bc = 0.

From the point of view of the analysis I am proposing, however, this kind of case is no longer a problem. I am now equating causal structure with the structure of equations with exogenous independence. And this analysis now gets Hesslow’s example right. The above equations capture the fact that B is a causal ancestor of T via two separate paths. That the coefficients conspire to stop this fact displaying itself in a correlation, as would normally happen, does nothing to invalidate the analysis. It just shows that the Faithfulness Condition is only a reliable rule of thumb, and not a necessary truth.

So the Faithfulness Condition now falls into its rightful place, as something that empirical researchers can generally rely on, but is in principle open to exceptions. There is nothing in the nature of causation to guarantee that probabilistic independencies should not arise by a cancelling out of parameters. This would be a freakish chance, but it is not built into the nature of causation. Unlike the Causal Markov Condition that takes us from correlations to causal links, the converse Faithfulness Condition that says that causal links display themselves in correlations is only delivered as a reliable rule of thumb.

Before proceeding, let me briefly observe that my account offers a natural explanation for the “asymmetry of overdetermination” invoked by analyses of causation in the tradition of David Lewis. This asymmetry consists in the fact that any time will contain many independent traces of past events, but scarcely any of future events. The Lewisian tradition assumes this asymmetry without further explanation[[9]](#footnote-9). But my account delivers it as a corollary of the way that, faithfulness failures aside, the joint effects of any cause will be correlated with each other. This means that the effects of any cause will tend to occur in concert, in a way its causes will not. And this in itself account for the fact that we have many traces of the past but not the future.

10 Single-Case Causation

The current suggestion is that C causes E if and only if it is an ancestor of it in a recursive system of structural equations with exogenous independence. This now promises a better hold on cases of single-case causation. (Did Joe Blogg’s smoking actually cause his cancer?)

Still, there are complications here. A natural first thought is that we can now say that some particular fact Ca is an actual cause of Ea if the recursive equations relating C and E imply that, in the actual circumstances, Ea depended on Ca. But then we have to address the issue, familiar from David Lewis’s attempt at a counterfactual analysis of actual causation, that there are cases of actual causation without counterfactual dependence, due to pre-emption, trumping, and so forth.

Fortunately, we can appeal here to wealth of recent work which uses “causal models” to address this issue. These causal models posit directed relationships, standardly portrayed by arrows, between actual and possible values of variables displayed by particular situations. The existing literature then aims to formulate recipes that will allow us to read off from the models which events some given result was actually caused by or counterfactually dependent on.

The analysis of this paper complements this literature. While much progress has been made on the way causal models can help analyse actual causation and counterfactuals, it is very unclear what features of the real world these models answer to. In particular, there is no agreed account of what the arrows in the models represent. To a large extent these relationships of asymmetric causal dependence are simply taken as given. The present paper offers a way of filling this lacuna. It suggests that we should take there to be an arrow between two variables just in case just in case one is a parent of the other in a system of deterministic equations with exogenous probabilistic independence.

The analysis of single-case counterfactual dependence in terms of causal models is relatively straightforward. Ea counterfactually depends on Ca if not-Ea follows from the equations plus not-Ca when we assign their actual values to all variables that are not downstream from C.

The recipe for reading off actual causation from a causal model is more contested. I myself am attracted to an analysis, due to Brad Weslake (forthcoming), which in effect argues that Ca actually caused Ea if the model shows that it was actually part of an INUS condition for part of an INUS condition for . . . Ea (where INUS conditionship is the familiar notion introduce by J.L. Mackie).

11 “C causes E”

It will be illuminating at this point to return briefly to the question of the meaning of generic causal claims of the form “C causes E“ (“smoking causes cancer”, “birth control pills causes thromboses”, . . .). I am now reading claims of this form as saying that C is an ancestor of E in a recursive system of structural equations. And I have shown how we can infer such claims, reliably if not infallibly, from correlations by using the bridge principles.

We might well wonder why we are so interested in claims of this form, given that they will standardly leave us in the dark about single-case dependencies. As observed earlier, knowing that smoking causes cancer won’t decide whether Joe Blogg’s cancer was actually caused by his smoking, or even whether it counterfactually depended on it. To know these things we’d need to about all the other equation-ancestors of cancer and what values they had in Joe Blogg’s case, and just knowing that “smoking causes cancer” is likely to leave us very much in the dark about this.

The answer is that claims like “smoking causes cancer” can be an informative guide to action even if they are fated to be uninformative about single-case relations. It is relevant here that generic “C causes E” claims will generally come with numbers attached. They will indicate *how much more probable* C rather than not-C renders E. And this of course is just what we need to know when we are deciding whether to do C in pursuit of E.

This issue deserves far more analysis that I can give here, but at first pass we can understand quantitative generic causal claims as showing, on weighted average across the contexts consistent with the agent’s knowledge of their situation, how probable it is that doing C will bring about a sufficient condition for E and not for not-E. Viewed like this, quantitative generic causal claims can thus tell agents how much reason they have to do C if they want E.

This now casts an extra light on the meaning of generic causal claims. I am currently committed to reading “C causes E” as saying that C is an ancestor of E in a system of generic casual equations. But not all such cases will be ones in which C on average makes a *positive* probabilistic difference to E. It might be an ancestor and yet make a negative probabilistic difference to E, or indeed make a zero probabilistic difference, as in faithfulness failures.

I am happy to accept that everyday usage might not be happy to say that “C causes E” when C has no net influence on E, as in failures of faithfulness, and moreover that everyday usage will downright resist “C causes E” when C lowers the probability of E. Still, I do no propose to pause on these terminological points. I have identified the crucial relation of ancestry in a system of deterministic equations with exogenous independence, and shown how this relation can be reliably though not infallibly evidenced by correlational patterns. I have then indicated in outline how this relation can be used to analyse a number of further relations, including actual causation, counterfactual dependence, and making a positive probabilistic difference on weighted average across a type of context. Showing how all these relations map onto everyday usage is a task for another time.

12 Quantum Mechanical Indeterminism

The reduction of causation I have proposed assumes that effects are always *determined* by antecedent facts. Values of dependent variables Xi are deterministic functions F(X1, . . . Xi-1, Ei) of the independent variables X1, . . . Xi-1, and the exogenous variables Ei. At first sight this might seem inconsistent with the indeterministic nature of the world revealed by quantum mechanics.

But let us not be too quick. Note that my reduction does not imply that everything is determined, only that *effects* are. Note also that it does not require that, at *every time* earlier than an effect, facts obtain that determine that effect, only that all effects be determined by facts that obtain *by the time* they occur.

This leaves it open that many of the facts that determine an effect might themselves be the outcome of quantum processes. The multiple influences that contribute to the exogenous variables could still be the outcomes of chancy quantum processes, and moreover the values of the exogenous variables might only become determinate shortly before the time of the relevant effect. That would be perfectly in line with the idea that the values of dependent variables are always deterministic functions of probabilistically independent exogenous variables.

Still, it doesn’t seem plausible to insist that all causal structures must be of this deterministic form. Suppose I place a bomb that is set to explode if a radioactive substance decays a certain amount in a certain interval. If the bomb explodes, then my action will have caused the explosion. But not every event in the causal chain from my action to the explosion will have been determined by the time it occurred. In particular, the relevant radioactive decay will have been a purely chancy matter.

We can expect many causal structures to share this form. Take the stock example of smoking and lung cancer again. Perhaps the causal route from smoking to cancer proceeds via the chancy breaking of certain bonds in DNA molecules. Again, this will mean that one of the steps in the causal chain running to smoking from cancer will not have been determined by the time it occurs.

At first pass, cases like these call for structural equations of a different form. In place of equations like

(19) E = aC + eE

we will need

(20) Chance(E) = aC + ePr(E)

This makes a significant difference to the analysis so far. This is because recursive structures of equations some of which only fix chances rather than definite values for their dependent variables are no longer guaranteed by the independence of their exogenous variables to satisfy the Markov Condition.

The best-known illustration of this is provided by the so-called Einstein-Podolsky-Rosen (“EPR”) correlations. In these cases, an initial state, together with further background factors including the setting of instruments, fixes chances for various quantum measurements made on two wings of an experiment involving spatially separated particles. But even if the background factors are probabilistically independent, the correlations observed between the measurement outcomes cannot be screened off by any feature of the initial parent state.

The relevant point here is that the non-determinism of the relevant equations *leaves room* for the outcomes to co-ordinate themselves within the freedom, so to speak, left open by their non-determination. And the EPR correlation shows that, once this room is made available, nature sometimes makes use of it.

Still, it is noteworthy that correlations with this non-Markov nature are unknown outside the physics laboratory. It requires very carefully arranged experimental circumstances to display the characteristic features of the EPR correlations. I take it that this is due to linked particles that might display EPR correlations quickly interacting outside experimental settings with macroscopic systems that are not designed to amplify the values of entangled variables. Because of this, the observable macroscopic events that are influenced by linked microscopic quantum systems will generally vary independently once we hold fixed any common sources of those linked events.

And this will then mean that structural equations in which chancy quantum events play a role will effectively be reducible to deterministic structural equations after all. For we can now rewrite equations of the form

(20) Chance(E) = aC + ePr(E)

as

(21) E = aC + ePr(E) + eE

where eE is a sort of “dummy variable” representing the way in which the chance of E resolves itself into actuality. And as long as we are dealing with cases, unlike the carefully arranged EPR set-up, where different chancy variables will so resolve themselves independently, we can assume that these extra chance-realizing dummy variables will be probabilistically independent of each other and the other exogenous variables.

We can usefully describe equations like (21) as “pseudo-deterministic”. Dependencies that appear probabilistic only because they omit the totality of determining factors are often termed “pseudo-*in*deterministic”. But contrast, while equations like (21) have the appearance of determining the effect variable E, this conceals the way that the eE terms on the right-hand side are an expression of the fact that nothing determines E until it occurs.

Still, as long as these dummy chance-realizing variables in a system of structural equations are probabilistically independent, then the equations will be function just like a system of deterministic equations with exogenous independence. So I now propose the following adjusted analysis of causation to accommodate the involvement of chancy events in causal structures.

(22) X causes Y if and only if it is an ancestor of Y in a recursive structure of deterministic *or pseudo-deterministic* equations with independent exogenous terms.

Because the exogenous variables function similarly in both deterministic and pseudo-deterministic equations, this analysis will imply the same connection between causes and correlations as before. The Causal Markov Condition with its correlative Linkage Principles will be a deductive consequence of the analysis, while the Faithfulness Condition and the Unlinkage Principles can be expected to hold except in special cases of faithfulness failure.

13 EPR Complications

What about correlations of the EPR sort? They might rarely manifest themselves outside the laboratory, but they can be real enough, and so raise a number of points.

Let us start by agreeing that EPR correlations between two spatially separated outcomes are not a matter of one outcome causing the other. This point perhaps deserves further discussion, but there are a number of prior reasons for accepting this, including the symmetry of the relationship between the two wings, and the impossibility of controlling results on one wing by manipulating the other.

Given this, EPR correlations are then counterexamples to the Causal Markov Condition. They do not signify that one outcome is causing the other, nor can they be viewed as due to a common cause of the outcomes consistently with the Causal Markov Condition, given that the correlations are not screened off by any feature of the initial state.

Now, as before, this is no problem for practical non-experimental researchers. As I have observed, we can be confident that we will not meet any observable EPR correlations outside the laboratory setting. So practical researchers can continue to assume the Causal Markov Condition and its Linkage corollaries in inferring causes from correlations.

Still, one might wonder where the EPR correlations leave my claim that the Causal Markov Condition is a deductive consequence of my proposed analysis of causation. If my analysis does indeed imply the Causal Markov Condition, and the EPR correlation show the Causal Markov Condition is not generally true, then that looks bad for my proposed analysis.

A crucial point here, however, is that the derivation of the Causal Markov condition from my analysis implied that variables in *causal structures* will satisfy the Causal Markov Condition, not that all variables whatsoever will—where causal structures means specifically recursive structures of equations with exogenous independence. The EPR correlations are not covered by this result, since, as we have seen, the equations governing the outcomes on the two wings in the EPR set-up cannot be put into the form of pseudo-deterministic equations with probabilistically independent exogenous variables.

This now means that my original Conditional and Unconditional Linkage Principles (2) and (3) were too generally formulated. They specified that certain causal implications follow from *any* correlations between variables. We can now see that this was too ambitious. The EPR correlations show us that the relevant casual implications are only guaranteed if we are dealing with variables which are governed by deterministic or pseudo-indeterministic equations with independent exogenous variables. This qualification to the Linkage Principles might be of no practical importance, given that no EPR-type correlations ever present themselves to non-experimental researchers, but it is needed if we want to keep the logic straight.[[10]](#footnote-10)

14 Conclusion

I have argued that the causal structure of the macroscopic world arises from the way macroscopic variables are governed by deterministic or pseudo-deterministic structures of equations with exogenous independence. Each macroscopic variable is a function of other macroscopic variables plus probabilistically independent exogenous variables. Because of this, any macroscopic effect can be attributed a sequence of probabilistically independent exogenous influences.

This is of course intended as an a posteriori reduction of causation, not as any kind of conceptual analysis. Some will be inclined to object to my account that they can conceive of a world in which the impact of one billiard ball alone deterministically causes the motion of another, without the help of any noisy background variables of the kind I say are needed to constitute causation. My response is that they may be able to conceive this, but what they are conceiving is impossible. (Papineau 1988.) Causation in the actual world, and in particular its directionality, metaphysically derives from probabilistic requirements on multiple causes, and would be absent from a world that lacked such complexity.

The story I have told is of course incomplete. It raises as many questions as it answers. In particular I have done nothing to explain why, whenever we find causation constituted by patterns of dependency with exogenous independence, the independent exogenous variables should turn out always to precede their dependent variables in time. But this itself is something that calls for explanation in a world governed by a temporally symmetric fundamental dynamics.

I take it that the answer is to do with the emergence of macroscopic phenomena from the quantum world. As indicated in the last two sections, it is natural to view the independence of exogenous variables as arising from the way that chancy quantum states manifest themselves in the observable macroscopic world. Now, different theories of quantum mechanics offer different accounts of the mechanics of quantum state “collapses”. Still, it is a constraint on these accounts that they should explain why such manifest collapses appear to occur asymmetrically in time. The story I have told is thus in a position to piggy-back on whichever theory gives the best account of apparent quantum state collapses.

References

Beebee, H. and Menzies, P. 2020 "Counterfactual Theories of Causation" in Zalta, E. ed *The Stanford Encyclopedia of Philosophy*(Summer 2020 Edition)

Blalock, H. 1964 *Causal Inferences in Nonexperimental Research* New York: W. W. Norton

Blanchard, T. and Schaffer, J. 2017 “Cause without Default” in Beebee, H., Hitchcock, C. and Price, H. eds Making a Difference, Oxford: Oxford University Press 175–214

Briggs, R. 2012 “Interventionist Counterfactuals” *Philosophical Studies* 160: 139–66.

Cartwright, N. 1979 “Causal Laws and Effective Strategies” Noûs, 13: 419–437

Cartwright, N. 1989 *Nature’s Capacities and their Measurement* Oxford: Oxford University Press

Cartwright, N. 2002 “Against Modularity, the Causal Markov Condition, and any Link between the Two: Comments on Hausman and Woodward” *The British Journal for the Philosophy of Science* 53**:** 411-53

Dearden, L. Ferri, J and Meghir, C. 2002 “The Effect of School Quality on Educational Attainment and Wages” *Review of Economics and Statistics* 4: 1-20

Dupré, J. 1984 “Probabilistic Causality Emancipated” in French, P., Uehling, T. and Wettstein, H. eds Midwest Studies in Philosophy IX Minneapolis: University of Minnesota Press 169–75

Durkheim, E. 1897 Suicide: A Study in Sociology translated by Spaulding, J. and Simpson, G. London: Routledge & Kegan Paul

Einstein, A., Podolsky, B. and Rosen, N. 1935 “Can Quantum-Mechanical Description of Physical Reality be Considered Complete?” Physical Review 47: 777–780

Elga, A. 2000 “Statistical Mechanics and the Asymmetry of Counterfactual Dependence” *Philosophy of Science* 68, 313-24

Fisher, R. A. 1925 *Statistical Methods for Research Workers* Edinburgh: Oliver and Boyd

Galles, D. and Pearl, J. 1998 “An Axiomatic Characterization of Causal Counterfactuals” *Foundations of Science* 3: 151–82

Good, I. J. 1961-2 “A Causal Calculus I–II” *British Journal for the Philosophy of Science* 11: 305–18, 12: 43–51

Glymour, C. 2004 “*Making Things Happen* by James Woodward” *British Journal for the Philosophy of Science*

Halpern, J. 2016 *Actual Causality* Cambridge Mass: MIT Press

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

Hausman, D. and Woodward, J. 1999 “Independence, Invariance and the Causal Markov Condition” *The British Journal for the Philosophy of Science* 50**:** 521-83

Healey, R. 1997 "Nonlocality and the Aharonov-Bohm Effect" *Philosophy of Science* 64: 18-41

Hesslow, G. 1976 "Two Notes on the Probabilistic Approach to Causality"*Philosophy of Science* 43: 290 – 92

Hitchcock, C. 2001 “The Intransitivity of Causation Revealed in Equations and Graphs” *Journal of Philosophy*, 98: 273–99

Hofer-Szabó G., Rédei, M., and Szabó, L. 2013 *The Principle of Common Cause* Cambridge: Cambridge University Press

Hoover, K. 2003 “Nonstationary Time Series, Cointegration, and the Principle of the Common Cause” *The British Journal for Philosophy of Science* 54: 527–51

Lazarsfeld, P. and Rosenberg, M. 1955. *The Language of Social Research: A Reader in the Methodology of the Social Sciences* Glencoe Ill: Free Press

Lewis, D. 1979 “Counterfactual Dependence and Time’s Arrow” *Nous* 13: 455-76

Loewer, B. 2007 “Counterfactuals and the Second Law” in Price, H. and Corry, R. eds *Causation, Physics, and the Constitution of Reality: Russell's Republic Revisited* New York: Oxford University Press, pp. 293-326

Papineau, D. 1988 "Response to Ehring’s ‘Papineau on Causal Asymmetry’” *British Journal for the Philosophy of Science* 4: 521-5

Papineau, D. 2001 “Metaphysics over Methodology” in Galavotti, M-C, Suppes,P and Costantini. D eds *Stochastic Causality*

Pearl, J. 2017 “The Eight Pillars of Causal Wisdom” edited transcript of a lecture at West Coast Experiments April 2017 <https://ftp.cs.ucla.edu/pub/stat_ser/wce-2017.pdf>

Pearl, J. 2018 *The Book of Why* London: Allen Lane

Pearl, J. 2000 Causality: Models, Reasoning, and Inference Cambridge: Cambridge University Press

Peters, J., Janzing, D., and Schölkopf, B. 2017 *Elements of Causal Inference: Foundations and Learning Algorithms* Cambridge Mass: MIT Press

Pollock, S. 2014 “Econometrics: A Historical Guide for the Uninitiated” *Working Paper 14/05* Department of Economics, University of Leicester

Reichenbach, H. 1956 *The Direction of Time* Los Angeles: University of California Press

Salmon, W. 1984 Scientific Explanation and the Causal Structure of the World Princeton: Princeton University Press

Schulz, K., 2011 “‘If You’d Wiggled A, then B Would’ve Changed’: Causality and Counterfactual Conditionals” *Synthese* 179: 239–51

Schurz, G., 2017, “Interactive Causes: Revising the Markov Condition” Philosophy of Science 84: 456–79

Schurz, G. and Gebharter, A. 2016 “Causality as a Theoretical Concept” *Synthese*

Simon, H. 1953 “Causal Ordering and Identifiability” in Hood, W. and Koopmans, T. eds *Studies in Econometric Method: Cowles Commission for Research in Economics*

Sober, E. 2001 “Venetian Sea Levels, British Bread Prices, and the Principle of the Common

Cause” *The British Journal of Philosophy of Science* 52: 331–46

Spirtes, P., Glymour, C., and Scheines, R. 1993 *Causation, Prediction and Search* Springer-Verlag New York

Spohn, W. 2001 "Bayesian Nets Are All There Is to Causal Dependence" in Galavotti, M-C, Suppes,P and Costantini. D eds *Stochastic Causality*

Suppes, P. 1970 A Probabilistic Theory of Causality Amsterdam: North-Holland Publishing Company

Wenglinksky, H. 2007 *Are Private High Schools Better Academically than Public High Schools?* Washington DC: Center for Education Policy

Weslake, B. forthcoming “A Partial Theory of Actual Causation” *British Journal for the Philosophy of Science*

Woodward, J. 2003 *Making Things Happen*

Woodward, J. 2016 "Causation and Manipulability" Stanford Encyclopedia of Philosophy https://plato.stanford.edu/archives/win2016/entries/causation-mani/

Wright, S. 1921 "Correlation and Causation" *Journal of Agricultural Research* 20: 557–85

Zhang, J. and Spirtes P. 2014 “Choice of Units and the Causal Markov Condition” in Guo G. and Liu C. eds Scientif*ic Explanation and Methodology of Science* Singapore: World Scientific 240-51

1. “Variable” can be understood as referring to a symbol on paper or in some other medium, to a function with abstract numbers as values used to model some worldly quantity, or to the worldly quantities themselves. My focus throughout this paper will be on the last-mentioned worldly quantities. Similarly, I shall understand “equation” as standing for worldly relationships between such quantities, and not for any symbolic or abstract numerical representation thereof. [↑](#footnote-ref-1)
2. Lewis original explanation was given in his “Counterfactual Dependence and Time’s Arrow”. Elga (2000) showed that Lewis’s treatment was insufficiently sensitive to the asymmetry of thermodynamics. Loewer (2007) remedied this deficiency but simply assumed without any further analysis that the asymmetry of overdetermination (”the predominance of local macro signatures of the past (but not of the future)” 317) is built into the asymmetry of thermodynamics. [↑](#footnote-ref-2)
3. By “correlation” I shall mean any case of non-independent probability distributions of two variables (school type/examination results, pollution/death rate, . . .) instantiated by some type of individual (schoolchildren, towns, . . .). I shall assume throughout that these are lawlike population patterns, as opposed to finite sample associations, and that they hold within some background field. [↑](#footnote-ref-3)
4. The more familiar coinage is directed acyclic “graph” (DAG). I have adopted “structure*”* instead to stress that my concern is with worldly relationships between worldly quantities, and not with the means by which we might represent these relationships. In such a causal DAS, an arrow means that the variable at the head of the is a cause of the variable at the tail. The arrows in such a DAS are required to be acyclic in the sense that that a variable can only be a causal ancestor of another if it is not also a descendant of it. Throughout this paper I shall assume that variables never reciprocally cause each other. When some coarse-grained variables seem to leave this as a possibility—for example, might not *happiness* cause *health*, and *health* also cause *happiness*?—then weshould switch to time-lagged versions of these variables, as in healtht1, healtht2, healtht3, . . . [↑](#footnote-ref-4)
5. The Causal Markov Condition says:

   In any directed acyclic structure of causal relationships, any variable will be probabilistically independent of every other variable (apart from its own causal descendants) conditional on its causal parents. (Cf Spirtes et al 1993 54.)

   A “structure of causal relationships” should be understood to include any set of causal relationships abstracted from reality. The Causal Markov Condition is only plausible if such structures are further understood to require that no common causes of included variables be omitted (for reasons elaborated in the next section). So understood, and supposing there are no further requirements on causal structures beyond these (but see footnote 10 below), Linkage (2) follows because any two causally unlinked variables can feature as parentless in a causal structure, and so must be uncorrelated, while Condition Linkage (3) follows because, in the absence of any links between the two variables that don’t involve {X}, controlling for {X} would screen off the correlation. [↑](#footnote-ref-5)
6. The Faithfulness Condition can be stated as:

   There are *no more* unconditional and conditional independencies than are required by the Causal Markov Condition. (Cf Spirtes et al 1993 56 [↑](#footnote-ref-6)
7. Let us now assume that our variables, including school type S, can be measured on some quantitative scale, for example by level of school funding. I shall also simplify by assuming throughout that all variables are measured from their means. [↑](#footnote-ref-7)
8. I take no view on the nature of lawlike deterministic connections in this paper. Everything I say is consistent with all the standard accounts of nomological necessity. [↑](#footnote-ref-8)
9. Loewer (2007) does aim to explain, in terms of the “past hypothesis”, why we have “records” of the past but not the future. But his account fails to explain why we have many separate such records, which is what he needs to account for the direction of causation. [↑](#footnote-ref-9)
10. When we assumed back in footnote 5 that nothing is required of causal structures beyond including all common causes of included variables, then the Causal Markov Condition implied the original Linkage Principles (2) and (3) in full generality. We need the restriction of causal structures to systems of equations to with exogenous independence to stop the Causal Markov Condition wrongly applying to quantum correlations. [↑](#footnote-ref-10)