Causation: One Word; Many Things

I am going to describe here a 3-year project on causality underway at LSE, funded by the

British Arts and Humanities Research Board. The central idea behind my contribution to the

project is Elizabeth Anscombe's (Anscombe 1971). I thus share a lot in common with the work

of Peter Machamer, Lindley Darden and Carl Craver, which is also discussed at these

meetings. My basic point of view is adumbrated in my 1999 book *The Dappled World*:

> *The Dappled World* takes its title from a poem by Gerard Manley Hopkins.
>
> Hopkins was a follower of Duns Scotus. So too am I. I stress the particular
>
> over the universal and what is plotted and pieced over what lies in one
>
> gigantic plane […]

> About causation I argue […] there is a great variety of different kinds of
>
> causes and that even causes of the same kind can operate in different ways
>
> […]

The term 'cause' is highly unspecific. It commits us to nothing about the

kind of causality involved nor about how the causes operate. Recognizing

this should make us more cautious about investing in the quest for universal

methods for causal inference. (Cartwright 1999, Ch. 5)

The defense of these claims proceeds in three stages.

Stage 1: As a start I shall outline troubles we face in taking any of the dominant accounts now on

offer as providing *universal* accounts of causal laws.[1]

1) The probabilistic theory of causality (Patrick Suppes) and consequent Bayes-nets methods of

    causal inference (Wolfgang Spohn, Judea Pearl, Clark Glymour)

2) Modularity accounts (Pearl, James Woodward, economist Stephen LeRoy)

3) The invariance account (Woodward, economist/philosopher Kevin Hoover)

4) Natural experiments (Herbert Simon, Nancy Cartwright)

5) Causal process theories (Wesley Salmon, Phil Dowe)

6) The efficacy account (Hoover)

---

[1] I exclude the counterfactual analysis of causation from consideration here because it is most plausibly offered as
an account of singular causation. At any rate, difficulties that the account faces are well-known.

Stage 2. If there is no universal account of causality to be given, what licenses the word 'cause' in a law? The answer I shall offer is: Thick causal concepts.

Stage 3. So what good is the word 'cause'? Answer: That depends on the assumptions we make in using it – hence the importance of formalization.

PART 1. The first stage is the longest. It involves a review of what I think are currently the most dominant accounts of causal laws that connect with practical methods. Let us just look at a few of these cases to get a sense of the kinds of things that go wrong for them. What I want to notice is a general feature of the difficulties each faces. Each account is offered with its own paradigm of a causal system and each works fairly well for its own paradigm. This a considerable achievement – often philosophical criticism of a proposed analysis points out that the analysis does not even succeed in describing the very system offered as an exemplar. But what generally fails in the current accounts of causality on offer is that they do not succeed in treating the exemplars employed in alternative accounts.

Bayes-nets methods. These methods do not apply where

- Positive and negative effects of a single factor cancel
- Factors can follow the same time trend without being causally linked

- Probabilistic causes produce products and by-products

- Populations are over stratified (e.g. they are homogeneous with respect to a common effect of two factors not otherwise causally linked)

- Populations with different casual structures or (even slightly) different probability measures are mixed

- …

- …

- … [2]

I will add one further note to this list. Recall that the causal Markov condition, which is violated in many of the circumstances in my list, is central to Bayes-nets. Advocates of Bayes-nets methods for causal inference often claim in their favor that "[a]n instance of the Causal Markov assumption is the foundation of the theory of randomized experiments" (Spirtes *et al*. 1996, 3).

But this cannot be true. The arguments that justify randomized experiments do not suppose the Causal Markov condition; and the method works without the assumption that the populations under study satisfy the condition. Using only some weaker assumptions that Bayes-nets

---

[2] For further discussion see Cartwright (2001a).

methods also presuppose, we can prove that an ideal randomized experiment will give correct results for typical situations where the causal Markov condition fails: e.g. cases of overstratification, the probabilistic production of products and by-products, or mixing.

Modularity accounts. These require that each law describe a 'mechanism' for the effect, a mechanism that can vary independently of the law for any other effect. I am going to dwell on this case because it provides a nice illustration of my general thesis.

So far I have only seen discussions of modularity with respect to systems like this:

$$x_1 \; c= u_1$$

$$x_2 \; c= f_2(x_1) + u_2$$

$$x_3 \; c= f_3(x_1, x_2) + u_3$$

.

.

.

$$x_m \, c= f_m(x_1, ..., x_{m-1}) + u_m {}^3$$

---

[3] The symbol 'c=' means that the left-hand and right-hand-sides are equal and that the factors on the right-hand-side are a full set of causes of the factor represented on the left.

where these are supposed to be causal laws for a set of quantities represented by V $= \{x_1, ..., x_m\}$ and where

    i) $\forall i \left(u_i \notin V \right)$

    ii) $\forall i \neg \exists j \, (x_i \, c \rightarrow u_j, x_j \in V)^4$.

Modularity requires that it is possible either to vary one law and only one law or that each exogenous variable can vary independently of each other. So

---

4 '$c\ c \rightarrow e$' means 'c is a cause of e'.

Modularity → either

(i) $x_1 \text{ c}= u_1$

$x_2 \text{ c}= f_2(x_1) + u_2$

$x_3 \text{ c}= f_3(x_1, x_2) + u_3$

.

.   .

$x_n \text{ c}= f_n(x_1, \ldots, x_{n-1}) + u_n \qquad \rightarrow \qquad x_n \text{ c}= X$

$x_{n+1} \text{ c}= f_{n+1}(x_1, \ldots, x_n) + u_{n+1}$

.

.

.

or

(ii)   there are no cross-restraints among the values of the u's.

Why should systems of causal laws behave like this? Woodward's main thesis is that this kind of modularity is the (single best) marker of what it is for a set of relationships to be causal.[5] He supports this with a lot of examples, but the issues he raises are frequently ones of identifiability, which are relevant only to the epistemology of causal laws, not to their metaphysics.

Hausman (1998) also takes modularity as central to the idea of causation. He adds an empirical consideration to support that systems of causal laws will always be modular. Although we may tend to focus on one or two or a handful of salient causal factors, in reality the cause of any factor is always very complex. This makes it likely that any two factors will always have some components of their total cause that are unrelated to each other and that thus can be used to manipulate the two factors independently. This may be plausible in the case of singular causation (with respect to purely counterfactual manipulations) that occurs outside any regimented system, but it does not seem true in systems where the causal behavior is repeatable and the causal laws depend on a single underlying structure.

I shall illustrate this below. But first I would like to look in some detail at an argument in support of modularity that has received less attention in the philosophical literature. Judea Pearl and Stephen LeRoy (1985) both make claims about ambiguity that I also find an echo of in Woodward. Causal analysis, Pearl tells us, 'deals with changes' – and here he

---

[5] See for instance (Woodward 1997 and 2000).

means changes under an 'intervention' that changes *only* the cause (and anything that must change in train (Pearl 2000, 345 and also Pearl 2002). So

> *Pearl/LeRoy requirement*: A causal law for the effect of x_c on x_e is supposed to state unambiguously what difference a unit change of x_c (by 'intervention') will make on x_e.

I always find it puzzling why we should think that a law for the effect of e on c should tell us what happens to e when the set of laws is itself allowed to alter or even when c is brought about in various ways. I would have thought that if there was an answer to the question, it would be contained in some other general facts – like the facts about the underlying structure that gives rise to the laws and that permits certain kinds of changes in earlier variables. My reconstruction of Pearl and LeRoy's answer to my puzzle takes them to be making a very specific claim about what a causal law is (in the kind of deterministic frameworks we have been considering):

A causal law about the effect of x_n on any other variable is Nature's instruction for determining what happens when either

1. The causal law describing the causes of x_n varies from $x_n = f_n(x_1,\ldots,x_{n-1}) + u_n$ to

   $x_n = X$

9

or

2. The exogeneous variable for x_n (i.e., u_n) varies *AND* **nothing else** varies except what these variations compel.

So for every system of causal laws:

- such variation in any cause must be possible, and
- the law in question must yield an unambiguous answer for what happens to the effect under such variation in a cause.

Hence the requirement called 'modularity'. But there must be something wrong with this conception of causal laws. When Pearl talked about this recently at LSE he illustrated this requirement with a Boolean input-output diagram for a circuit. In it, not only could the entire input for each variable be changed independently of that for each other; so too could each Boolean component of that input. But most arrangements we study are not like that. They are rather like a toaster or a laser or a carburetor.

I shall illustrate with a casual account of the carburetor, or rather, a small part of the operation of the carburetor – the control of the amount of gas that enters the chamber before ignition. I take my account of the carburetor from Macaulay's book (1988). Macaulay's account is entirely verbal (and this shall be important to my philosophical point later on). From the verbal account we can construct the diagrammatic form that the functional laws governing the amount of gas in the chamber must take:

1. gas in chamber c= f (airflow; α) pumped gas + (α') gas exiting emulsion tube

2. airflow c= g (air pressure in chamber; β)

3. gas exiting emulsion tube c= h (gas in emulsion tube, air pressure in chamber; γ)

4. air pressure in chamber c= j(suck of the pistons, setting of throttle valve; σ)

$\alpha = \alpha$ (geometry of chamber, …)

$\alpha' = \alpha'$ (geometry of chamber, …)

$\beta = \beta$ (geometry of chamber, …)

$\gamma = \gamma$ (geometry of chamber, …)

$\sigma = \sigma$ (geometry of chamber, …)

Look at equation 1. The gas in the chamber is the result of the pumped gas and the gas exiting the emulsion tube. How much each contributes is fixed by other factors. For the pumped gas

both the amount of airflow and a parameter α, which is partly determined by the geometry of the chamber; and for the gas exiting the emulsion tube, by a parameter α', which also depends on the geometry of the chamber. The point is this. In Pearl's circuit-board, there is one distinct physical mechanism to underwrite each distinct causal connection. But that is incredibly wasteful of space and materials, which matter for the carburetor. One of the central tricks for an engineer in designing a carburetor is to ensure that one and the same physical design—for example, the design of the venturi—can underwrite or ensure a number of different causal connections we need all at once.

Just look back at my diagrammatic equations, where we can see a large number of laws all of which depend on the same physical features—the geometry of the carburetor. So no one of these laws can be changed on its own. To change any one requires a redesign of the carburetor, which will change the others in train. By design the different causal laws are harnessed together and cannot be changed singly. So modularity fails.[6]

My conclusion though is not that we must discard modularity. Rather it is not a universal characteristic of some univocal concept of (generic) causation. There are different causal questions we can ask. We can, for instance, ask the causal question we see in the Pearl/LeRoy requirement: how much will the effect change for a unit change in the cause if the unit change in the cause were to be introduced 'by intervention'? The question will make sense and have

---

[6] For more details see Cartwright (2001b).

an unambiguous answer for modular systems. The fact that many systems are not modular does not mean that this is foolish question to ask when systems are modular.

Woodward's invariance account. This is a strengthening of the modularity account. Modularity accounts tell us that causal laws predict what happens under variations of the appropriate sort. Woodward's *invariance account* says that if a claim predicts what happens under variations of the appropriate sort, it is a causal law. Hence some of the problems for this claim:

- Invariance works only for systems that are modular, not for toasters and carburetors
- I can prove Woodward's invariance claims (once formulated explicitly) *for special systems*. Among the axioms for these systems are
  - ➢ Transitivity
  - ➢ Functional Dependence
  - ➢ Anti-symmetry and irreflexivity
  - ➢ Uniqueness of coefficients
  - ➢ Consistency
  - ➢ No functional relations obtain not derivable from causal laws.[7]

---

[7] For full axioms, see Cartwright forthcoming.

This last forbids, e.g., that two variables might show the same time trend. So invariance also has its special problems.

But there is one thing to note in favor of invariance methods—unlike Bayes-nets methods, they can give decisive answers about specific causal hypotheses even where the causal Markov condition fails. For instance, this is true for linear probabilistic structures like those below, where the u's serve to introduce genuine irreducible probabilities.

$$x_1 \; c = b_1 + u_1$$

$$x_2 \; c = a_{21}x_1 + u_2$$

.

.

.

$$x_n \; c = \sum_{i=1}^{n-1} a_{ni}x_i + u_n$$

In any case in which the u's are not mutually independent, the causal Markov condition will not hold. Nevertheless invariance methods will give correct judgments about individual causal hypotheses. That is, correctly formulated invariance methods will work even when the u's are correlated leading to violations of the Markov condition. On the other hand, because we need variations of just the right sort, where the 'right sort' is specified in causal terms, invariance

methods require a great deal more specific antecedent causal knowledge than do Bayes-nets methods. Hence they are frequently of little use to us.

Natural experiments. If we want to tie method – really reliable method – and "analysis" as closely as possible, probably the most natural thing would be to reconstruct our account of causality from the experimental methods we use to find out about causes.[8] Any such attempt is bound to illustrate my overall point. The conditions that must obtain for a situation to mimic that of an experiment are enormously special. A notion of causality geared to conditions that obtain in an experimental setting – whether it occurs naturally or is contrived by us – is not likely to fit well for a large variety of commonly occurring systems that other accounts (and ordinary intuitions as well) will count as causal.[9]

Causal processes. These accounts require that there be a continuous space-time process that conveys the causal influence from cause to effect. There is a large literature looking at the problems that arise for various specific versions of the account. But, as Kevin Hoover argues, none of them will work for crucial cases in economics that we want to study: say, cases of equilibrium, where causes and effects are 'simultaneous'; or cases involving causal relations among quantities all of which only make sense when measured over extended periods of time

---

[8] See Simon (1953)
[9] For a further discussion see Cartwright (2003).

– which may well then overlap with each other. Hoover himself offers an account which can deal with such cases.

Hoover's effective strategies account. [10] 'X c→Y' if anything we do to affect X will affect Y as well, but not the reverse, maintains economist/methodologist Kevin Hoover. But Hoover's characterization is too weak to serve as a universal condition on what it means for x to cause y. Consider the following pattern which we might see in a mechanical device like the toaster, where I draw the causal arrows in accord with our primitive intuitions about how the device operates – intuitions that will probably also be in accord with a causal process account of causal laws. In this case Hoover allows that *x causes y*, so long as u and v are factors that can be directly manipulated.

---

[10] See Hoover (2001). There are also a number of well-argued "agency" accounts in the philosophical literature. I focus on Hoover's because it is ties most closely with methodology, which is the central interest I have in finding an adequate account of causality. Also, I imagine Hoover's version of an agency account will be less familiar to philosophers of science and my discussion can provide an introduction to it.
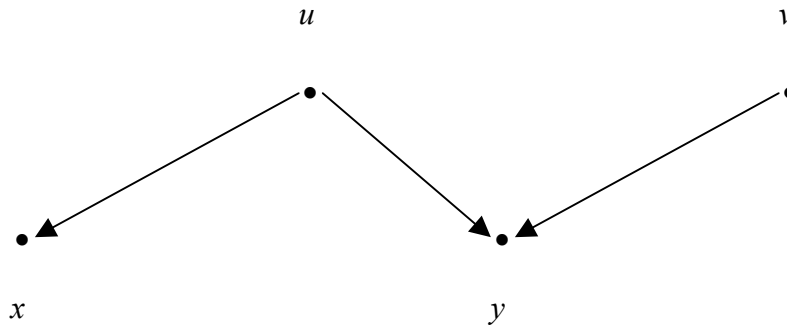
u   v

x   y

Figure 1

So Hoover's condition is too weak.

On the other hand it is also too strong, since it never allows that *x* causes *y* or the reverse when the association between the two is given thus (again the arrows represent causal process causality or perhaps probabilistic causation):
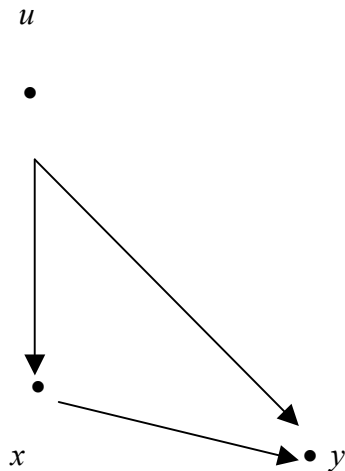
u

•

↓ ↘

• → • y
x

Figure 2

Hence Hoover's account is too strong. Nevertheless it is based on a causal question whose answer may matter enormously to us: can we affect *y* by affecting *x*?

Diagnosis. All these accounts have problems. Does that mean that none of them are any good and we should throw them out? To the contrary, I think they all are very good. They fail, I hypothesize, because the task they set themselves cannot be accomplished. Under the influence of Hume and Kant we think of causation as a single monolithic concept. But that is a mistake. The problem is not that there are no such things as causal laws; the world is rife with them. The problem is rather that there is no single thing of much detail that they all have in

common, something they share that makes them all *causal* laws. These investigations support a two-fold conclusion

- There is a variety of different kinds of causal laws that operate in a variety of different ways and a variety of different kinds of causal questions that we can ask

- Each of these can have its own characteristic markers; but there are no interesting features that they all share in common.

PART 2. All the accounts I described seem to suppose that there is one thing – one characteristic feature – that makes a law a *causal* law. I want to offer an alternative. Just as there is an untold variety of quantities that can be involved in laws, so too there is an untold variety of causal relations. Nature is rife with very specific causal laws involving these causal relations, laws that we represent most immediately using content-rich causal verbs: the pistons *compress* the air in the carburetor chamber, the sun *attracts* the planets, the loss of skill among long-term unemployed workers *discourages* firms from opening new jobs… These are genuine facts, but more concrete than those reported in claims that use only the abstract vocabulary of 'cause' and 'prevent'. If we overlook this, we will lose a vast amount of

information that we otherwise possess: important, useful information that can help us with crucial questions of design and control.

To begin to see this alternative picture, consider again the causal equations above that describe the operation of an automobile carburetor. Where did this equation schema come from? As I said, I constructed them from the description of the carburetor in *How Things Work.* If you look there you will find a far more content-rich causal theory about carburetors than could be represented in equations like the ones I propose, even when the functional forms are all filled in properly. Here are some of the more specific laws that are represented by my set of causal equations. (Of course, in an engineering treatment the laws would be both quantitative and more detailed.)
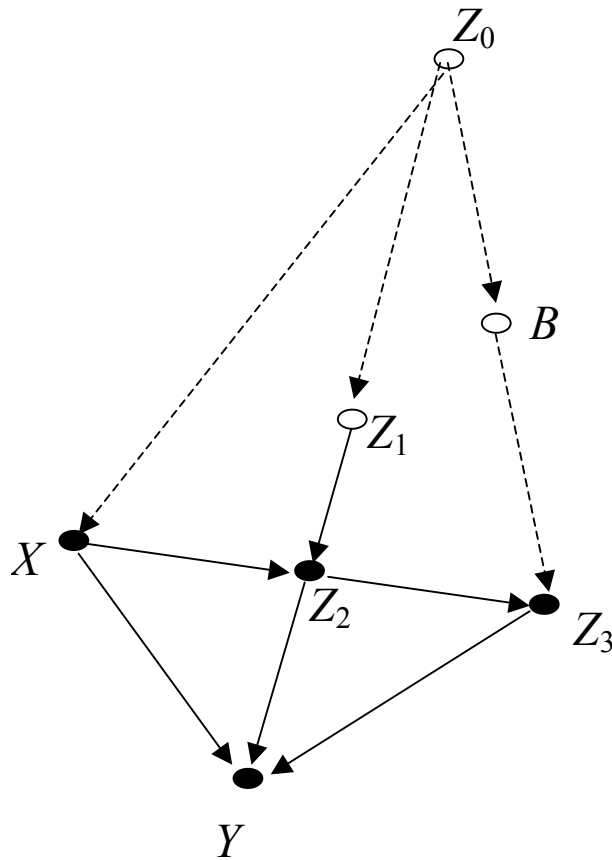
- the carburetor *feeds* gasoline and air to a car's engine…
- the pistons *suck* air in though the venturi…
- the low-pressure air *sucks* gasoline out of a nozzle…
- the throttle valve *allows* air to flow through the nozzle…
- pressing the pedal *opens* the throttle valve more, *speeding* the airflow and *sucking in* more gasoline…
- …
- …
- …

These law claims express details of the laws that govern the operation of the carburetor that are missing from the equations. If there is any doubt, just consider all the things one can learn from these kinds of thick nomological descriptions that one cannot learn from the equations. For instance, suppose we wish to increase the acceleration produced by stepping on the accelerator and we think of doing so by increasing the width of the venturi (thus allowing more gas through). Our attempt will probably be counterproductive because doing so will also affect the drop in pressure in the air as it passes through and thereby the amount of gas that can be sucked out of the nozzle.

For a Bayes-nets example, consider a case that Judea Pearl often discusses, "an experiment in which soil fumigants (X) are used to increase oat crop yields (Y) by controlling the eelworm population (Z) but may also have direct effects, both beneficial and adverse, on yields beside the control of eelworms…farmer's choice of treatment depends on last year's eelworm population (Z_0)…the quantities Z_1, Z_2, and Z_3 represent, respectively, the eelworm population, both size and type, before treatment, and at the end of the season…B, the population of birds and other predators" (Pearl 1995, 669).

Here is the Bayes-net diagram that Pearl offers to represent the situation he describes (Pearl 1995):

Figure 3: A causal diagram representing the effect of fumigants, *X*, on yields, *Y*. Source: Pearl 1995, p. 670.

X: fumigants
Y: yields
B: the population of birds and other predators
Z_0:   last year's eelworm population
Z_1:   eelworm population before treatment
Z_2:   eelworm population after treatment
Z_3:   eelworm population at the end of the season

It is clear that we could give a thicker description of the causal laws operating in this experiment. Perhaps the soil fumigant *poisons* the infant eelworms, or perhaps it smothers the eelworm eggs, or…; and any of a vast number of activities could be covered by the claim that the soil fumigant has independent beneficial or adverse effects on yields. Perhaps the fumigant *enriches* the soil or *clogs* the roots. Instead Pearl gives an even thinner description. He replaces all the thick descriptions by one single piece of notation—the arrow. The arrow represents in one fell swoop all the different causal-law relations described in the thicker theory.

There is one important fact to note about thick causal concepts: they are not themselves composites from a noncausal law and some further special characteristics that make it a causal law—e.g. characteristics of the kind I have just been reviewing. Consider a comparison. Just as I contrast general causal terms like *cause* and *prevent* with thicker ones like *compress* and *attract* and *smother*, Bernard Williams in *Ethics and the Limits of Philosophy* contrasts general evaluative terms like *good* and *ought* with "'thicker' or more specific ethical notions […] such as *treachery* and *promise* and *brutality* and *courage*, which seem to express a union of fact and value" (Williams 1985, 129).

But, Williams explains, they only *seem to* express a union of fact and value: these terms are not composites made up of two parts, a description with an evaluation added on. Elsewhere I give a whole set of arguments about causation that exactly parallels Williams's about ethical concepts (Cartwright 2002). Here I note only one significant point. All thick causal concepts imply 'cause'. They also imply a number of noncausal facts. But this does *not* mean that 'cause' + the noncausal claims + (perhaps) something else implies the thick concept. For instance we can admit that *compressing* implies *causing + x,* but that does not ensure that *causing + x + y* implies compressing for some non-circular *y.*

PART 3. I have presented the proposal that there are untold numbers of causal laws, all most directly represented using thick causal concepts, each with its own peculiar truth makers; and there is no single truth maker that they all share in virtue of which they are labeled 'causal' laws. What job then does the label 'causal' do?

When it comes to formal systems, we can say a lot about what job it does. That is the beauty of the formal system. The idea is that whether it is right to call something by the general term *cause* or not, depends on what you are going to do with that label once you have attached it. Consider Pearl's work. *If* your causal relations, described by thick causal concepts, satisfy Pearl's modularity assumptions, he shows you a wealth of counterfactual conclusions,

predictions about results of manipulations and techniques for corroboration of specific hypotheses that you are entitled to make about these relations.

Or consider my formalizations of different versions of Woodward's invariance claims. *If* the Cartwright axioms are all satisfied for a given set of thick causal concepts, we can prove that an observed functional relation among quantities corresponds to a true causal claim iff the relation provides correct predictions under the right variations.

We can further prove things like the following:

- A system of true causal-law claims including y c= $\sum a_i x_i + u_i$ will make correct predictions about y if any of the causes of any of the x_i anywhere back in the chain are varied in the right way.

- Suppose we add assumptions that guarantee that there is a chain of causal laws between x_i and y. Then it is easy to show that if, for all i, any of the intervening factors between x_i and y vary 'to zero' in the appropriate way, y will no longer depend on x.

I also think analogous things are true even when scientific theories or claims will not bear formal reconstruction. There is still a loose set of inferences fixed by the context to which we are entitled when we make a causal-law claim with the thin word 'cause' in it. The correctness of the term 'cause' will depend on the soundness of the conclusions we draw.

To summarize. Formalisms using thin causal concepts can be very useful. They provide conditions that thick causal laws might satisfy, conditions that license a specific body of inferences. General schemata using thin causal concepts are crucial for scientific practice. For they provide us with ready-made methods. Otherwise we have to find the appropriate method for each new system of laws we confront.

But there is no guarantee that we have, or can readily construct, formal schemata that will fit every system of laws we encounter. The causal arrangements of the world may be indefinitely variable. We may after all live in a dappled world.

REFERENCES

Anscombe, E. (1971), *Causality and Determination*, Cambridge: CUP. Reprinted in Sosa and Tooley (1993, 88-104).

Cartwright, Nancy (forthcoming), "Two Theorems on Invariance and Causality", *Philosophy of Science.*

Cartwright, Nancy. (2003), "How to Get Causes from Probabilities (ala Simon)", LSE MS.

Cartwright, Nancy (2002), "Causation, Thick and Thin", University of Nottingham Lecture and LSE MS.

Cartwright, Nancy (2001a), "What is Wrong with Bayes Nets?" *The Monist*, 84:2 , 242-264.

Cartwright, Nancy (2001b), "Modularity: It Can – and Generally Does – Fail", *Stochastic Dependence and Causality,* D. Constantini, M.C. Galivotti, P. Suppes (eds), Stanford: CSLI Publications, 65-84.

Cartwright, Nancy (2000), "Measuring Causes: Invariance, Modularity and the Causal Markov Condition"*, Measurement in Physics and Economics Discussion Paper Series Monograph DP MEAS 10/00*, London: Centre for Philosophy of Natural and Social Science.

Cooley, T. and LeRoy, S. (1985), "Atheoretical Macroeconometrics: A Critique", *Journal of Monetary Economics*, 16:3, 283-308.

Hoover, Kevin (2001), *Causality in Macroeconomics*, Cambridge: Cambridge University Press, ch. 6.

Hausman, Dan (1998), *Causal Asymmetries,* New York: Cambridge University Press.

Macaulay, David (1998), *The Way Things Work*, Kindersley Dorsey Publishing.

Pearl, Judea (2002), "Causal Modelling and the Logic of Science", LSE Lakatos lecture.

Pearl, Judea (2000), *Causality: Models, Reasoning and Inference*. Cambridge: Cambridge University Press.

Pearl, Judea (1995), "Causal Diagrams and Empirical Research", *Biometrica* 82, 669-710.

Herbert A. Simon (1953): Causal Ordering and Identifiability. In W.C. Hood and T.C. Koopmans (eds.): Studies in Econometric Method. New York: Wiley, 49-74.

Sosa, E. and Tooley, M. (eds) (1993), *Causation*. Oxford: Oxford University Press.

Spirtes, P., Meek, C., and Richardson, T. (1996), "Causal Inference in the Presence of Latent

Variables and Selection Bias", *Technical Report CMU-77-Phil*.

Williams, B. (1985), *Ethics and the Limits of Philosophy*, Cambridge, Mass.: Harvard University Press.

Woodward, J. (2000), "Explanation and Invariance in the Special Sciences", *BJPS* 51, 197-254.

Woodward, J. (1997), "Explanation, Invariance and Intervention", *PSA 1996*, vol 2, S26-41.