

Durham Research Online

Deposited in DRO:

21 September 2015

Version of attached file:

Accepted Version

Peer-review status of attached file:

Peer-reviewed

Citation for published item:

Cartwright, N. and Efstathiou, S. (2011) 'Hunting causes and using them : is there no bridge from here to there?', *International studies in the philosophy of science.*, 25 (3). pp. 223-241.

Further information on publisher's website:

<http://dx.doi.org/10.1080/02698595.2011.605245>

Publisher's copyright statement:

This is an Accepted Manuscript of an article published by Taylor Francis Group in *International Studies in the Philosophy of Science* on 30/09/2011, available online at: <http://www.tandfonline.com/10.1080/02698595.2011.605245>.

Additional information:

Use policy

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a [link](#) is made to the metadata record in DRO
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the [full DRO policy](#) for further details.

*Hunting Causes and Using Them:
Is There No Bridge from Here to There?*¹

Nancy Cartwright
with Sophia Efstathiou

LSE and UCSD

Abstract

Causation is in trouble –at least as it is pictured in current theories in philosophy and in economics as well, where causation is also once again in fashion. In both disciplines the accounts of causality on offer are either modelled too closely on one or another favoured method for *hunting* causes or on assumptions about the *uses* to which causal knowledge can be put – generally for predicting the results of our efforts to change the world. The first kind of account supplies no reason to think that causal knowledge, as it is pictured, is of any use; the second supplies no reason to think our best methods will be reliable for establishing causal knowledge. So, if these accounts are all there is to be had, how do we get from method to use? Of what use is knowledge of causal laws that we work so hard to obtain?

I. Two Actions: Hunting Causes and Using Causes

Philosophic – and economic – accounts of causality are almost always rooted in ideas about either how to HUNT causes or how to USE them;

¹ This paper was presented at the Society for Philosophy of Science in Practice (SPSP) meeting in London in the summer of 2007. It develops ideas presented in a more roundabout way in Cartwright (2006) ‘Where is the Theory in our “Theories” of Causality?’ *Journal of Philosophy*, Vol. CIII, no. 2. 2006, pp. 55-66 and in Cartwright (2007) *Hunting Causes and Using Them*, Cambridge University Press. Support for the research and completion of the paper was provided by the AHRC project *Contingency and Dissent in Science*, by the University of California at San Diego Senate, by the University of California Humanities Research Institute, by the Institute for Advanced Study at Durham University and by the Spencer Foundation; we are grateful to all for their help.

but not both simultaneously. Most are almost directly read off either a favoured method of establishing causal claims – putting the label ‘causal’ onto a claim – or a favoured use we expect to make of causal claims – inferences to claims of practical use. Accounts based on methods for hunting causes include the probabilistic theory of causality and its descendent Bayes-nets theories, accounts based on experimental methods (like the Galilean experiments discussed below), most versions of causal process theories and theories that rely on exchanges of conserved quantities or the like and the usual counterfactual accounts. Those rooted in use include manipulation and intervention accounts, those based in causal decision theory and counterfactual accounts that allow realistic implementations.

Almost all these accounts are good ones, Cartwright has argued – good for specific purposes in specific kinds of systems.² And, taken together, they cover both rules for entry into the language of causality and for exit from causal language. The problem is that there is no bridge from one to the other. What assures us that the knowledge we hunt at such great effort and cost can be put to the uses we want to make of it? To be practicable a theory of causation must simultaneously ground how we label features as ‘causes’ and the inferences we make once the label is attached. So we need a theory of causation that gives us in one fell swoop both methods for inferring causes and methods for using them.

For the purposes of this paper, trying to remain relatively neutral about just what causal laws are since we will consider a variety of different accounts, we propose that one think of a causal law as a law-like causal regularity: ‘C causes E’ constitutes a causal law if C regularly causes E – and that’s no accident.

To illustrate our points about hunting and use we shall discuss four philosophical accounts of causal laws besides Cartwright’s own

² Nancy Cartwright (2007), *Hunting Causes and Using Them*, Cambridge University Press, Chapter 2, pp. 11-24.

- James Woodward’s ‘level’ invariance account³
- related accounts based on ‘Galilean’ experiments⁴
- Lewis-style (‘miracle’-based) counterfactuals⁵
- probabilistic theories of causality.⁶

All these are geared towards hunting causes: they ensure that we are “correct” in putting causal labels on. But they say little about what we can do with “causal” knowledge once we have it. On the other side we shall describe Woodward’s ‘modularity’ assumption and the kind of assumptions economists are wont to make about causality that guarantee its usefulness.⁷ We pick these for discussion because they allow us to illustrate our points fairly simply and vividly.

II. Two Problems: Unstable Enablers and External Validity

We tend to assume that the knowledge we secure with our best methods for testing casual laws carried out in the best circumstances is knowledge that we can use directly: in knowing the causal law we know how to change effects by changing their causes and we can make reliable predictions about the results of so doing. We seem to take some connection there as given.⁸ But why should one make this assumption? When if ever is it justified?

³ James Woodward (2003) *Making Things Happen*, Oxford University Press, New York.

⁴ Cf Ronald Giere’s *Understanding Scientific Reasoning*. New York: Holt, Rinehart & Winston, 1979.

⁵ David Lewis (1973) “Causation”, *Journal of Philosophy*, 70, 556-67.

⁶ Esp. Patrick Suppes (1970) *A Probabilistic Theory of Causality*, Amsterdam: North Holland.

⁷ Among philosophers we discuss Woodward rather than other accounts (eg. Peter Menzies and Huw Price (1993) “Causation as a secondary quality”. *British Journal for the Philosophy of Science* **44**, 187—203, Douglas Gasking (1955) “Causation and Recipes”, *Mind* **64**, 479-87 or Georg Henrik von Wright (1971) *Explanation and Understanding*) because his two separate conditions directly illustrate the distinction between characterizations of causality rooted in hunting methodologies and those that work to guarantee use.

⁸ Cf. Cartwright’s early paper ‘Causal Laws and Effective Strategies’, *Nous*, 13, 419-37 (also published in Cartwright (1983) *How the Laws of Physics Lie*, Oxford University Press) or the spate of work at about the same time on causal decision theory (for instance in Harper, W. L., Stalnaker, R. and Pearse, G. (eds.) 1981, *Ifs: Conditionals, Belief, Decision, Chance and Time*, Dordrecht: Reidel.)

Philosophic accounts of causation based on hunting methodologies give no satisfactory answer to these questions. The conditions our accounts need to secure causal knowledge are not sufficient to secure the inferences that would put them to use. These accounts face two problems.

1. *Unstable enablers*: changes in the enabling factors that support causal laws – a problem commonly recognized in economics.
2. *External validity*: the problem of generalizing from a particular setting or population to the one of interest.

II.1 Unstable Enablers

Economists have worried about the stability of causal laws and their use for policy prediction since at least the time of John Stuart Mill.⁹ We go to the trouble of setting economic policy on the assumption that the factors we propose to use as policy levers can change and that our causal laws will allow us to predict what will follow on from that. But features in the economy that permit change in the ways we like and can predict allow for change in ways that we don't like and worse, often in ways that we can't predict. Our very attempts to use our causal laws may undermine the laws themselves. Instabilities in the causal laws we establish may not only result when we interfere to set policy, they may arise naturally as well, and at times and places we often have no way of knowing. What use then are causal laws for prediction?

Mill was pessimistic. He argued that economics cannot be an inductive science because the background arrangements of causes fluctuate and do so erratically, both naturally and as a result of deliberate actions by us. Any effect consequent on a particular cause depends, he argued, on a large background of other causal factors simultaneously at work as well,

⁹ John Stuart Mill (1836 [1967]) "On the Definition of Political Economy and on the Method of Philosophical Investigation in that Science", reprinted in *Collected Works of John Stuart Mill*, vol. IV, Toronto: University of Toronto Press.

factors we are rarely able to identify. The effects consequent on a cause on one occasion, or over a period of time, cannot be relied on to occur on other occasions since this myriad of other causes is likely to change and in ways we cannot predict. So, concluded Mill, the sea of other causal factors that enable a particular factor to bring about a particular effect are rarely in place long enough to allow us to make new predictions based on past regular associations between the two factors.

Besides variations in the arrangements of causes themselves Mill was also concerned about a second kind of possibly unstable enabler: changes in the underpinning structures that give rise to causal laws in the first place. It is a common experience of everyday life that what causal laws hold in a situation depends on the underlying structure that gives rise to them. Pressing a lever in a toaster causes the bread to drop into the toaster and brown, pressing a lever on the floor of the driver's side of a Rover causes the car to accelerate. Putting a pound coin into a vending machine in the UK causes a bag of crisps to drop into the tray, putting a pound coin into a similar looking machine in the USA just gums up the works. For most of these man-made devices the usual ways of manipulating the cause to achieve the predicted effect will not gum up the works. They are typically well-shielded to prevent just that happening.

The structural features that give rise to causal laws at work in the economy may be far more porous however. Like Mill, the early econometricians worried about just that. More recently Chicago School economists like Robert Lucas argue more strongly: not only *can* active policy intervention change the underlying structural arrangements in ways that undermine the very causal laws that are being relied on to predict the outcomes of those interventions; it is very likely this *will* happen. If interventions are expected, so argues the Chicago School, agents will change their behaviours and in just such a way as to undermine the established causal relations that predictions are based on.¹⁰

¹⁰ Robert Lucas (1976) "Economic Policy Evaluation: A Critique", in Lucas, R. (1981) *Studies in Business Cycle Theory*. Basil Blackwell. Oxford, and Robert Lucas (1988) "On the Mechanics of Economic Development", *Journal of Monetary Economics*, January 1988, 22, 3-32.

On the other hand, despite being worrisome, causal laws can be unstable in predictable, even happy ways. Take an example from Mill's *On the Subjugation of Women*.¹¹ In this essay Mill notes that many (like August Comte, one of his chief adversaries on this issue) maintain that there is a well-established (very law-like) causal regularity between being a woman and being inept at leadership, reasoning and imagination. Happily this causal law is not stable under changes in the underlying enabling structure that supports it. Change the social structure so that the upbringing, roles, education and experiences of women are more like those of 19th century British middle-class men and the causal laws relating sex to leadership, reasoning and imagination shift dramatically. The enabling factors that hold the regularity fixed change in more or less expected ways, and luckily so.

To see how the problem of unstable enablers plays out in a current philosophical account of causation, let us turn to James Woodward.¹² Woodward claims that two requirements must be satisfied by a relation before we can call it a causal law. The first requirement is *level invariance* (*modularity* comes later):

- *Level invariance*: a relationship between putative causes and an effect is level invariant if the relationship stays fixed as any putative cause in the relationship varies 'by intervention'.

'Intervention' is hard to define properly. It is something like a 'miracle' in the *Lewis account of counterfactuals*: a change in the cause at the last stage while all other causes from the set of causal laws operating in the situation remain the same, except for changes induced by the change in the targeted factor (sometimes called by philosophers, changes 'causally downstream' from the targeted change).

¹¹ John Stuart Mill (1869 [1997]) *The Subjection of Women*, Dover Publications, 1997.

¹²Cf. James Woodward (2003) *Making Things Happen*, Oxford University Press, New York and Sandra Mitchell (2003) *Biological Complexity and Integrative Pluralism*, Cambridge University Press, Cambridge, who stresses the need for invariance without a detour through causality.

When it comes to causality Woodward focuses on systems of linear deterministic causal laws that look like this:

Linear deterministic causal law

$$z \text{ c= } ax + by$$

The symbol 'c=' means that the left and right-hand side are equal and that the factors on the right are a complete set of causes of those on the left. (Reference to the population and circumstances is repressed as is usual in presentation.) Applied to $z \text{ c= } ax + by$, *level invariance* dictates that the functional relationship $z = ax + by$ must hold as x varies¹³ while y and all the true causes in causal laws at work in the situation (except those causally downstream from the change in x) stay fixed; and the functional relation also holds as y varies and x and other causes stay fixed.

Woodward supports his claims arguing by contradiction: he considers various examples in which level invariance is violated by relations known to be spurious.¹⁴ The canonical example is the relationship between a storm and the reading of a barometer. Storms are caused by fluctuations in the relative pressure of atmospheric fluids so we can relate the presence of a storm to a function of atmospheric pressure by writing, say,

$$(1) \text{ storm} = f(\text{pressure}).$$

The level of a well-functioning barometer, an instrument designed to measure pressure, should also depend on atmospheric pressure, according to a relationship like

$$(2) \text{ barometer level} = g(\text{pressure}).$$

¹³ The range of permitted variation of variable must be specified as well. Mention of the relativization to these ranges is suppressed as well.

¹⁴Cf. James Woodward (2003) *Making Things Happen*, Oxford University Press, New York.

One can mathematically solve (2) for pressure in terms of the inverse function g' and substitute into (1) to relate the presence of a storm to the reading of a barometer:

$$(3) \text{ storm} = f(g'(\text{barometer level})).$$

The question is which of (1), (2) or (3) describes a genuine causal law.

The true causal laws at work in the situation contain three variables: pressure, storm, barometer level. We expect that equation (1) would continue to obtain if pressure were varied 'by miracle' with barometer level staying fixed. So it should pass the level invariance test. So too does (2) since it would not change if pressure were changed by a miracle while the presence or absence of the storm stays fixed. Equation (3) on the other hand, while it may be useful for calculating the likelihood of a storm given that the instrument is functioning correctly, does not express a causal law and this shows up in the level invariance test: if the pressure stays fixed while the barometer level changes 'by miracle' (say it just explodes), the relationship between the storm and the reading of the barometer changes. It is well-known that you can't bring on the storm by breaking the glass!

Cartwright has elsewhere supported Woodward's emphasis on level invariance by showing it that is a *sufficient* condition for causality.¹⁵ She does so by proving a kind of representation theorem for level invariance vis-à-vis causal laws. The theorem uses some fairly straightforward axioms that a set of causal laws should satisfy – like asymmetry, irreflexivity and the assumption that any functional relations that hold are generated by genuine causal relations. These are all axioms that are presupposed in most discussions of causality and in particular in the examples that Woodward employs. The theorem then shows that any functional relation generated by a set of causal laws will be one of those

¹⁵ See Cartwright (2003) 'Two Theorems on Invariance and Causality', *Philosophy of Science* 70, pp. 203-224; reprinted in Cartwright (2007) *Hunting Causes and Using Them*, Cambridge University Press.

causal laws if and only if it is level invariant.¹⁶ So given a set of ‘true’ causal laws like (1) and (2) no spurious relation derived from them will be level invariant.¹⁷

Despite the fact that level invariance ensures that a functional relationship is a causal law given the axioms laid down for causal laws, Woodward thinks this is not enough. He adds a second condition, *modularity*, which we shall discuss below. But Cartwright, to the contrary, takes the theorem to prove that level invariance is sufficient. In that case level invariance is a good – indeed sure – way to HUNT causes, to put a causal label onto a relationship. If a relationship satisfies this condition we have conclusive reason to think it is causal. But does level invariance get around Mill and Lucas’s problem?

Unfortunately it does not. (And we will later argue that modularity doesn’t really do better either.) Looking for level invariant laws presupposes that factors in causal laws are not erratically interacting with the causal structures that give rise to them.¹⁸

Consider again Mill’s original example of the subjugation of women. The relationship between sex and leadership could well have been constant across a very great many variations in other causes affecting leadership

¹⁶ Cartwright takes these axioms to be fairly innocuous and to be true of causal laws even if her singular-causings account of causal laws is mistaken. NC notes that there has been some objection that the axioms are not so innocuous because a transitivity axiom is included. But the transitivity axiom assumes only that if x appears as a cause of y in a linear deterministic causal system, we still have a causal law for y if we substitute for x the right-hand side of any causal law that has x as effect. This is necessary unless we are willing to assume that causation in nature is not continuous in time, so that there is a notion of direct causal law (the ‘last’ law in operation before the effect is produced) that is not representation relative and that it is this notion of direct causal law that we are trying to characterize.

¹⁷ibid.

¹⁸ It also supposes a relatively clear separation between the structure (what Cartwright calls the ‘nomological machine’) that gives rise to a set of causal laws and the laws themselves. Cartwright has defended this division in *Nature’s Capacities and their Measurement* (1989) OUP and in *The Dappled World* (1999) CUP. NC does not think it is universally applicable though. But she does think it can be made practically everywhere where there are causal laws at work that can be represented in the usual triangular array of equations, whether those equations represent deterministic or probabilistic causality and whether they are linear or not. For more discussion see Efstathiou (manuscript) “Nomological Machines In Scientific Practice”.

over the centuries, constant enough to count it as level invariant. It was throughout a true causal law: being a woman caused one to be weak at leadership. But if we transform the structure of society changing the role and experiences of women from birth onwards, the relationship may well no longer satisfy Woodward's condition of level invariance. So the well-established causal law would be no guide for predicting the effect of sex on leadership skills if the structure is changed. The invariance of a relationship under changes in variables in the other causal laws at work does not ensure predictability unless underlying enabling structures stay fixed. But that we cannot expect structural causal enablers to remain stable *is just the Lucas critique!*

II.2 External validity

A second problem many current theories of causality have to grapple with is *external validity*. This is a well-known problem in methodology. Many of the methods that allow us to establish causal-law claims most securely can be applied in only very narrow settings. We establish results very securely in a particular experimental setting or a particular test population. But the method itself provides no basis for extending the results to a population or setting different from that in the test. Theories of causality that are too closely tied to these kinds of method will suffer from the same problem; the very way causal laws are characterized makes causal laws very narrow in their scope and thus very limited in their predictive power.

Consider a standard method for hunting causes: *Galilean experiments*. The goal in a Galilean experiment is dual. First we wish to eliminate all confounders, then to establish a law-like regularity between cause and effect with no confounders to interfere. We can eliminate confounders by physically isolating an experimental system from background interference and/or by making various idealizing assumptions. The problem is that establishing the presence of a causal relation in absence of

confounders does not ensure that the same causal relation persists once confounders are present.

Suppose we establish via a perfect Galilean experiment that C causes E. Call the set of confounding factors – the ones we work so hard to eliminate in the experiment – ‘N’ (for ‘nasties’). Then what we have actually established in the experiment is

$$C \& \neg N \rightarrow E$$

from which we cannot infer

$$C \& N \rightarrow E.$$

But that is the kind of result we need for reliable prediction if we want to use C to control E in real life settings. Inferring that a cause will have the same effect whether or not confounders are present is a logical fallacy.

Any characterization of causal laws that reads them off from the results of a Galilean experiment will have even worse troubles. In that case causal laws will end up *by definition* to hold only in situations where confounders are absent.

External validity is also a problem for *David Lewis’s counterfactual* account of causation. *Lewis counterfactuals* establish a causal link by changing the cause at the last instant by a miracle-like intervention that removes the cause while leaving everything else the same. If the effect does not obtain once the cause – and only the cause – is absent then a causal relationship is established.

Although Lewis offers his account as an account of singular causal relations, it is also the basis for related accounts of causal laws. But as an account of causal laws it suffers from the problem of external validity. Any inferential power we get from *Lewis counterfactuals* pertains to a particular setting – whatever setting is under consideration, the setting in

which the miracle will intervene to change the putative cause. Even though the setting is not an idealized one, if we follow Lewis we establish a law-like causal regularity for only one kind of setting. Unless further assumptions are made there are no guarantees that the causal law established is resilient to a different background arrangement of confounders. So Lewis counterfactuals can lay down causal laws for settings where confounders occur, but only for the specific arrangement of them under consideration. It is no help for any different arrangement.

What we know from *Lewis counterfactuals* is that for some specific N

C&N $c \rightarrow E$.

This does not imply any results for some different N'. In particular it is a logical fallacy to conclude that

C&N' $c \rightarrow E$.

The same cause need not have the same effect under a new arrangement of confounders.

External validity is also a problem for an account more akin to Cartwright's own: the *probabilistic theory of causality* developed by Patrick Suppes. This account is almost read off statistical methods for inferring causal laws from observational (as opposed to experimental) data: stratify on all possible confounders, then look for correlations between the putative cause and effect within each stratum. One fairly good attempt at formulating the probabilistic theory (for yes-no variables) says that C is a cause of E *in a particular arrangement* of confounders, say K if and only if the cause increases the probability of the effect in that arrangement: $P(E/C\&K) > P(E/\neg C\&K)$.¹⁹ This of course

¹⁹ The formulation given here still isn't quite right because K must not hold fixed any causal intermediaries by which C causes E on a given occasion. Cartwright's own best attempt relies on reference to singular causings even in the formulation of the probabilistic theory. See Cartwright (1989) *Nature's Capacities and their Measurement*, Oxford University Press, sections 2 and 3 and

makes it explicit that casual laws are situation relative.²⁰ C causes E *simpliciter* if it does so in all arrangements of confounders.

But once more, the mere hap that under a particular arrangement of confounders the cause increases the probability of the effect gives little information about what happens in a different arrangement of confounders. As before

$$P(E/C\&K) > P(E/\neg C\&K)$$

does not imply

$$P(E/C\&K') > P(E/\neg C\&K').$$

We have again no logical grounds for inferences in settings different from the ones our laws are derived in. And again as before, if we use the probabilistic theory to characterize what a causal law is the problem is even worse since by definition the same law cannot hold in different settings. One may of course look to the non-relativized law: C causes E *simpliciter*. In this case causal laws become more useful but incredibly hard to hunt – we then need reason to suppose that the increase in probability will hold across *all* arrangements of confounders. That we certainly do not get just by looking at what happens in any one or handful of such arrangements.

So, here we see three familiar accounts of causality – the probabilistic theory, the counterfactual account, and accounts read off from the theory of the Galilean experiment. All three are rooted in methods for hunting causes. And all three are very good at what they do – putting a causal label on a relationship only where it seems correct to do so. That's not surprising since the methods they are based on are among our surest

Cartwright (2007) *Hunting Causes and Using Them*, Cambridge University Press, for a fuller discussion.

²⁰ As usual the population relativity is buried in the probability measure. The laws hold *for any population* for which the assumed probability measure holds.

methods for casual inference. But all three face the trouble of external validity once it comes to USING causes.

Galilean experiments establish conditions sufficient for attaching a causal label on a feature. If changes in C and C alone are followed by E in idealized settings with no confounders this is sufficient for C to be a cause of E. But notice, the changes must be in C alone and the setting must be ideal. Similarly *Lewis counterfactuals* are supposed to give conditions sufficient for attaching a causal label. If changes in C and C alone are followed by E given a specific arrangement of confounders this is sufficient for C to cause E – but only in cases where C alone is changed and where the confounders are arranged in just the same way. *Probabilistic causality* tells us to attach a causal label if changes in C and C alone are followed by an increase in the probability of the effect given a specific arrangement of confounders. This is again sufficient for C to cause E but only when C alone is changed in just the specified arrangement of confounders.

How then can we use causal knowledge obtained in a *Galilean experiment* for predictions about what will happen when confounders are in place or when factors other than C change as well? How can we infer something from one “nasty” setting about what effects follow the cause in another one using *Lewis counterfactuals*? How can we use *probabilistic causality* to infer from one case where the presence of the cause makes the effect more probable something about other cases where the same cause is known to be present but other causes have changed?

These are well known problems in the methodology of causal inference. But methodologists have a way to deal with them that philosophers cannot adopt. Methodologists often seem to suppose that there is something, they know not what exactly it is – a causal law – that can be established in a variety of *different* ways, then used for prediction in a variety of ways and situations different yet again. Causal laws are something like the charge of an electron or valency of oxygen in this respect.

We philosophers, however, aim to say what a causal law is, or at least to give some significant characterizing features of it, and to do so in a way that makes sense both of the ways we hunt causes and the ways we use them. But many of our attempts are too operationalistic so not surprisingly they suffer from the standard troubles of operationalism. If a concept is defined too closely to one or another way we test for it, we cannot account for why other test methods are reasonable and we cannot underwrite the usual inferences we make with the concept. Conversely, if a concept is defined too closely to one or another of the inferences we make using it, nothing about the concept supports other inferences nor grounds any methods for testing if the concept applies.

Current theoretical accounts of causation based on hunting methodologies are sufficient for identifying the cause of an effect in a particular situation; they do not give us methods for navigating to new settings. They only tell us where we are, not how we can get to where we want to go.

III. Unstable Enablers and External Validity: The Same Problem from Different Perspectives

Although the two problems of unstable enablers and external validity have a different source, when it comes to the structure of inference there is a sense in which they are equivalent. This suggests a deeper epistemological question that needs answering and also that there might possibly be a common answer to them both.

We worry about “unstable enablers” when we study a particular case – e.g. a set of agents, the direction of a particular effect, a particular variable – and we care to keep studying *this* case in midst of a – possibly – changing background structure. The problem of “unstable enablers” becomes visible after we have identified a causal relationship (e.g. pushing down the lever producing toast, putting the pound coin in the

slot producing a packet of crisps, the mass of the sun causing the earth to orbit around it, stepping on the throttle causing the car to accelerate) within some perceived structure; some possibly actual but for our purposes conceptualized space (like a toaster, the vending machine, the planetary system or a Rover 500 automobile). We take the case we are interested in as momentarily *fixed against this underlying structure* and worry about how the changing structure may affect what we care about. (e.g. What would happen if the toaster were wet? If the machine ran out of crisps?, etc.) The problem of “unstable enablers” speaks of shifts of known and unknown structural factors which seem to happen in *time* and affect the case we fixed our interest on.

Compare this with the problem of “external validity”. External validity becomes an issue when we shift our interest from a case already studied to one, hopefully similar, but *other* than the one studied. We think of the problem as arising when we export causal knowledge gained from a particular case to some new situation of interest. (e.g. Does pressing down the lever work in any toaster? Is the collision impact on the test dummy the impact suffered by the driver of a Rover 500?, etc.) The problem of “external” validity speaks of changes as arising when we move in *space*, from places “interior” to those “exterior” to our case study. So we might say that the problem becomes relevant when we think of causal knowledge as shifted across space rather than when things change with time.

Let’s take a step of abstraction. Using “space” and “time” variables, Σ and T , we could re-describe the two problems as follows:

1. Unstable Enablers:

Hypothesis & Structure (Σ_1, T_1) => Conclusion

but Hypothesis & Structure' (Σ_1, T_2) => *New Conclusion*

2. External Validity:

Hypothesis & Setting (Σ_1, T_1) => Conclusion

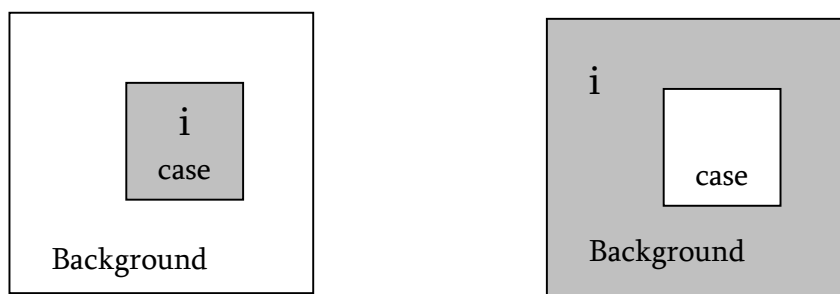
but Hypothesis & Setting' (Σ_2, T_1) => *New Conclusion*

The differences in the effect depend on change that occurs across a perceived temporal or a spatial setting:

1. Unstable Enablers: $\Delta E = \varphi(\Delta T)$
2. External Validity: $\Delta E = \psi(\Delta \Sigma)$

The “space” and “time” variables, Σ and T , here do not refer to physical space-time, but rather to dimensions of our experience (or better yet, our talk) of things causally interacting in physical space-time. But the way we talk about these problems is contingent. If we think of these dimensions as interchangeable, talking about here and there becomes the same as talking about now and then. When we take the instability of causal enablers as arising in time, before and after we modify policy say, we get what we have termed the problem of “unstable enablers”; when we take the instabilities of causal factors to occur across space – from the test population to a target population say – we get the problem of external validity. The two problems are equivalent in this sense (**Figure 1** depicts how the distinction is made).

Figure 1 What changes from one labelled problem to another is what holds our interest not the logical structure of the situation.



(1) **Unstable Enablers:** interest (i) fixed on a case; worry about how background affects the case studied

(2) **External Validity:** interest (i) shifted to a case “outside” the case studied; worry about whether our case applies here

So what? There is no denying that unstable enablers and external validity are different methodological problems for us. Well, two things are made visible here. First, that we need to address both problems when drawing up our methodology *at the same time*. The problem of external validity it

would seem is always present no matter (or on top of) whether we take the “underlying” causal structure to be changing or stable. Conversely, why should we think that the underlying structures are stable during the time it takes us to export our causal knowledge? Looking at the two problems from the perspective of changes in ‘space’ and ‘time’ brings home how likely it is that we face both problems much of the time.

Besides noting that the two problems are probably happening at the same time [for which one can argue without abstractions] our point is to note that the problems share a form and may also share a solution on a more abstract theoretical plane than either policy or methodology.²¹

IV. Cartwright’s Causal Laws and Capacities

Cartwright’s account of causal laws takes singular causation as primary and builds laws from there: C causes E in Φ if and only if some C ’s regularly cause E ’s in Φ (in the ‘long run’). Cartwright then argues that where probabilities apply, the probabilistic theory of causality as formulated in section II.2 provides necessary and sufficient conditions for causal laws. This is a good study to consider because it has all the problems we raise, writ large. As a version of the probabilistic theory, this account of causal laws is entirely rooted in a hunting methodology, a very reliable methodology if ideally carried through, but concomitantly very narrow in the range of the claims established and hence of extremely limited use. This is reflected in the fact that causal laws, as Cartwright sees them, are always relative; in her account they are relative both to a particular arrangement of confounders and to the nomological machine – the underlying structure – that gives rise to the causal regularity.²² This

²¹ For further discussion see Efstathiou (manuscript) “From Methods to Use: Is there a Here and There?”.

²² Cf Cartwright (1989), *Nature’s Capacities and their Measurement*, Oxford University Press. Also note as in section II.2 that we can drop the relativization to the arrangement of confounders by relativizing instead to a population. Clearly C will cause some E ’s in a population if it is guaranteed to cause some E ’s in some subpopulation of that population (ie. a subpopulation that is homogeneous with respect to confounding factors).

means Cartwright's causal laws are *fragile*: they are open to both the problems of external validity and of unstable enablers.

What then about use? Cartwright gets this by an altogether different route, not from causal laws at all but from capacities. Capacities are powers that agents possess to contribute in a fixed way to what results whenever they are present. For example, because of their nuclear structure permanent magnets have the capacity to attract metallic objects. This is a capacity we measure in various experiments but the experiments themselves are not enough to tell us that there is a capacity to be measured in the first place. The claim that magnets have such a capacity is grounded in a large extended and complicated network of theory, experiment and successful prediction.

Because the capacity will produce its contribution whenever it is present (or is properly triggered)²³ knowledge of capacities can be extremely useful. But beware. For the magnet, 'attraction' is the contribution, not the actual motion (or not) that occurs when the magnet operates. The attraction is always there even if the metallic object never moves. This is a piece of information we can use reliably across a huge variety of situations. But the predictions it gives rise to may not be as helpful as we wish. The metallic earring is stuck between the floorboards. Shall we buy a magnet to get it out? We can reliably predict that the magnet will attract the earring but we need a whole lot more information to predict that the earring will move. The most we can definitely predict with the kind of knowledge we usually have in these situations is that the magnet *may very well* pick up the earring that fell between the floorboards.

So, even with capacities, predictive power is *weak*. But that is not the point here. What matters for our worries about causal laws is that causal laws and capacities are entirely distinct. This is so even if one has a very different account of causal laws from Cartwright's. None of the accounts of causal laws currently discussed in philosophy look anything like an

²³ Except for chancy capacities, which produce their contributions only spasmodically.

account of capacities; nor do our standard methods for testing causal laws serve well for establishing capacities. Capacities can be of use for predicting what happens when we set policy or build new technological devices. That, however, does not salvage causal laws. In introducing capacities Cartwright never meant to undermine causal laws. But focussing on the distinction between capacities and causal laws points up the problem in bold relief: we are very good at finding out about causal laws; but once we have done so, of what possible use are they?

V. Solutions?

There are a number of strategies one might adopt to deal with worries about the usefulness of causal laws. None really work.

a. Chuck Causal Laws. The first, obvious solution to these problems is to chuck causal laws. Go for what you need. In Cartwright's case this is capacities. It is not a causal law that is in operation whenever the magnet succeeds in picking up the earring. It is instead the capacity. But Cartwright's account is metaphysically heavy. It postulates powers and in exactly the way Hume despised. Capacities require a three-fold distinction between 1) the presence of the capacity (e.g. whenever the magnet is present so too is the capacity to attract metallic objects); 2) the exercise of the capacity (the attracting of the earring) and 3) the actual result (the movement – or not – of the earring). Hume allows at most two of these but all three are necessary to do the job.

If we follow *Sandra Mitchell's* attack on 'laws', we can also chuck causality.²⁴ Mitchell points out that for use we do not need laws; ipso facto we do not need causal laws. *Any* truth can be useful so long as it is *true* where you propose to use it for prediction.

²⁴ Sandra Mitchell (2003) *Biological Complexity and Integrative Pluralism*, Cambridge University Press, stresses the need for invariance without a detour through causality.

Classic *instrumentalism* can also be seen as a way of chucking causality and going directly for what we need, much along the lines of the proposal we take from Mitchell. Science need not establish laws; what's needed are instruments that give us correct (enough) results for the predictions we want to make.

As liberating as it sounds, chucking causality doesn't get rid of our troubles. First of all it does not salvage causal laws. Are these just a wasted effort after all? Second, we still need to be told both how to establish and how to use whatever substitute notions are proposed. Which claims will be true in a particular setting, how should we use them and how do we establish that they will be true? Chucking causality gives no bridge from language entry to exit; from method to use. It just promises you can build another bridge, further down the road.

b. Dubbing. A second solution is to go for what you need and *dub* it causality. This approach is typical in econometrics. What makes for causality in econometrics is "structure", represented in structural equations. So, what's structure? Structural equations are more or less the ones that can be relied on for the predictions we want. But again this builds no bridge from language entry to exit; from method to use. Econometrics is very good at telling how to estimate parameters in structural equations – this looks much like language entry. But there's no good theory of what structure is that fits with both some good theory of 'model adequacy' – what makes it okay to assert a set of 'structural' equations in the face of data – and simultaneously with the assumption that the model can be used for prediction under intervention or in new situations.

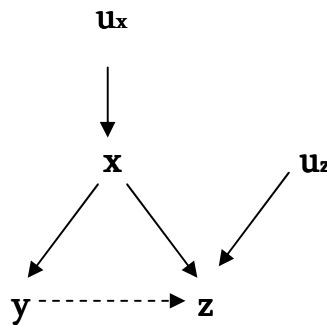
To reinforce this point, let us turn to an account of causality offered by macroeconomist and methodologist *Kevin Hoover* that seems geared very much to use, unlike the philosophical accounts we have looked at so far and more like familiar manipulation and intervention accounts in

philosophy.²⁵ (We discuss Hoover rather than the philosophers in part as a way of introducing his work to readers outside the philosophy of economics and partly because we can illustrate our point so cleanly with it.) Hoover defines causality directly in terms of the effects that can be achieved by manipulation, real manipulation.

*Hoover: C causes E iff anything we can do to fix C partially fixes E but not the reverse.*²⁶

Although this definition secures a connection between causation and manipulation the kinds of relations it calls ‘causal’ would not count as causal in everybody’s books – like probabilistic theories of causation, causal process theories or Lewis-style counterfactual accounts. **Figure 2** provides an example of a simple mechanism to illustrate, where the *u*’s are ‘policy levers’ – quantities we can manipulate, and the solid lines with arrows depict pure ‘mechanical’ causation, like pushing on a lever at one end to trip a switch at the other. The dotted line depicts ‘Hoover’ causation.

Figure 2



²⁵ See Kevin Hoover (2001) *Causality in Macroeconomics*, Cambridge University Press, Cambridge. It should be noted that the description given here of Hoover’s account is not one he is happy with. Cartwright claims it is what his definitions say and take the kind of causal relation described by the definitions as a very important one different from more ‘mechanical’ kinds of causal relations. He maintains that he intends his account to cover the more conventional notion of ‘mechanical causation’ and that various caveats he offers allow his definitions to do so. For further discussion, see Cartwright (2007) *Hunting Causes and Using Them*, Cambridge University Press, Chapter 14.

²⁶ Hoover causation thus is closely associated with the kind of ‘implementation neutral’ counterfactual that Daniel Hausman proposes for investigating casual claims, but with the range of implementations restricted to implementations we are able to bring about. See Cartwright (2007) *Hunting Causes and Using Them*, Cambridge University Press, Chapter 16.

If Hoover²⁷ is right about what causation is, we can see why causes by their very nature can provide predictions about strategies for manipulating the world. But many will not want to allow that Hoover's is an account of causation at all. More important, the account is entirely rooted in use and does not supply any bridge to get there from method. Knowing Hoover-causal facts will certainly be helpful in predicting the effects of policy interventions. But how does one learn these facts in the first place? Indeed many standard methods for causal inference will yield wrong verdicts vis-à-vis Hoover causation (e.g. methods that look for causal processes and physical connections), many will often not be applicable (e.g. Galilean experiments and Lewis counterfactual investigation in cases where the 'miracle'-like intervention is not among those we can do), and many will give 'no' answers to causal relations where a more nuanced verdict would be of far more help (e.g. where E changes under some manipulations of C but not others, C will not Hoover-cause E).

c. *Add-ons*. A third way is to be more demanding. Require conditions for both method and use. Woodward takes this route. Besides level invariance he requires '*modularity*' for a relationship to pass as a causal law.

Modularity: there must be at least one way to 'intervene' to change any genuine cause.²⁸

The effect of this requirement is that each variable appearing in a causal law operating in a situation can be changed without changing anything else except the effects of changing that variable. This again is a 'miracle'-like change. What justifies this as a condition on causality? Woodward is

²⁷ Or better, with footnote 25 in mind, 'Hoover as described here'.

²⁸ This is not exactly how Woodward defines modularity but it is how he uses the notion sometimes and especially to do just the job discussed here. See Woodward's definition of modularity in Woodward (2003) *Making Things Happen*, Oxford University Press, New York, p. 329.

clear:²⁹ this addition allows us to use the relation in question for predictions about manipulations. That, we take it, is why Woodward calls his account of causality indifferently an ‘invariance’ account and a ‘manipulability’ account.

As an answer to our worries it is not satisfactory however. First, it underwrites the use of causal laws for predicting the outcomes of manipulations not for the manipulations we might be envisaging, but only for very special kinds of miracle-like interventions that change the cause and nothing else. These are the kinds of manipulations that are demanded in a Galilean experiment or in Lewis-style counterfactuals. That may well be how they come to play such a special role in Woodward’s account. They are good for a very special way of testing for causal laws. But they are no good for showing why knowledge of causal laws is useful for real policy predictions.

In this respect econometricians David Hendry and Robert Engle do better.³⁰ First they require that causes be *exogenous*.³¹ This is a technical notion that has to do with efficient estimation from data. That facilitates language entry. What about language exit? They require that as well: before they will call a relation ‘causal’ they demand that it not only be exogenous but also ‘superexogenous’, which is a relative notion. A relation is *superexogenous* relative to an envisaged policy change just in case it will remain true under that change.

Accounts like Woodward’s and Hendry and Engle’s are what we call ‘add-on’ accounts. They do not provide an account, a theory, of causality

²⁹ Actually, he gives the same reason – causes must be usable to manipulate their effects – for both level invariance and for modularity. We cite it only for modularity because level invariance does not provide manipulability unless modularity is added and Cartwright at any rate has an alternative defense of level invariance.

³⁰ See Hendry (2001) *Causality in Macroeconomics*, Cambridge University Press, Cambridge and Hendry (2004) ‘Causality and Exogeneity in Non-stationary Economic Time-Series’, *Causality: Metaphysics and Methods* Technical Report CTR 18-04, CPNSS, London School of Economics and Engle, R. , Hendry, D. and Richard, J. F. (1983), “Exogeneity”, *Econometrica*, 51, 277-304.

³¹ Also, in line with the probabilistic theory of causality there is in general the assumption that causes and effects are probabilistically dependent.

rich enough to justify our usual rules for both language entry and language exit for causal laws. In particular they offer nothing that shows why our standard methods for inferring causal laws leads us to claims that can be used in the ways proposed. They just refuse to call something a causal law unless it can *both* be admitted by their favourite test for causality AND can be relied on for the kinds of predictions they describe.

This is no bridge at all from method to use. It is mere mereology. And mereology will not save causal laws. Without the assurance of a bridge from standard method to use, the assumption that the claim can be used in the ways described needs to be established on its own, independently. But then why bother with the first half to begin with? What's the point of all the time, money, thought and effort that goes into putting the causal label on through careful use of our best methods if the tests that do all the work in justifying our prediction are altogether different ones? If mereology is the only alternative left, it seems preferable to give up on causality altogether and instead adopt one of the first two strategies: give up on causal laws and establish what we need. Then call it causality if you like

VI. Aside: Woodward versus Cartwright

It might be useful before concluding to juxtapose Cartwright's capacities and Woodward's modularity demand. Woodward's conditions on causality concern *whole equations*:

$$z = ax + by$$

If an equation is *level invariant* under miracle-like interventions on right-hand-side variables, it must be a causal law under Cartwright's axioms. *Modularity* demands that we call it a causal law only if there is a miracle-like intervention for each right-hand-side variable.

Cartwright's capacities are causal tendencies associated with individual features. The strength of capacities is measured in Galilean experiments.

In an equation like the one above, if the quantities designated by x and y both have capacities with respect to z , then 'a' is the strength of x 's capacity to produce z and 'b' is the strength of y 's capacity to produce z . Note that there is no assumption that the equation is invariant. What equation holds is relative to a particular setting, a particular arrangement of causes. The coefficients however are invariant. The hypothesis that x and y have capacities with respect to z guarantees that they make the *same* contribution (*viz* ax and by) whenever they are present. So they can be used to build new equations to describe situations where they are present with other causes and without each other.

Woodward's mechanisms and Cartwright's capacities are both difficult to establish. Beyond that though they have complementary limitations and advantages. For Woodward, the whole causal equation is invariant. This is good for prediction but bad for scope; we are restricted to predictions in new situations that have exactly the same set of causes operating. Further, Woodward causal laws describe what happens under miracle-like interventions. This is again bad for scope. We cannot easily perform miracles so our causal laws won't help much with predictions about real life changes. Still, for Woodward there is at least *a* way to use a causal law for prediction and maybe we'd better just find that way.

Capacities on the contrary are good for scope. Once established capacities can be carried to new settings (addressing worries of external validity) and even, for 'fundamental' capacities, across different underlying structures (dealing with unstable enablers).³² For example, now that it has been established that magnets have the capacity to attract metallic objects, the attraction may be confidently relied on in new settings. But capacities are not as good as we might hope for prediction. What is guaranteed with a capacity is that it will produce a fixed *contribution*. That's the bit that Hume would not like. What actually happens is far

³² A good many capacities are derivative however. These too will depend on the underlying structure, or 'nomological machine', that gives rise to them.

harder to predict since it depends on what other causes are operating and what all their contributions together add up to.

There are nice cases of course, in the sciences, where we know a set of capacities, each with its fixed contribution, and we also know a rule of composition for how the contributions 'add'.³³ Forces in mechanics and their rule of vector addition is the paradigm. And it is this paradigm that Mill turns to when he defends the role of stable tendencies (from which Cartwright's capacities are copied) in the economy. But as Anna Alexandrova and Julian Reiss argue, both from their studies of economic models,³⁴ for most cases, even cases where we are strongly inclined to ascribe capacities, Mill's hope for a rule of composition is daft.

Consider for example the case of the subjugation of women. We might well admit that women do have the capacity for leadership and intelligence. What will result when this capacity operates in a variety of real-life settings? Is it really reasonable to assume that in each setting, a set, albeit possibly a large one, of other causes each with its own capacity is at work and that the outcome of all acting together can be calculated by a fixed rule of composition? Even in the case of the magnet this picture seems suspicious. To be sure, there are cases where all the causes affecting the motion of a metallic object can be represented neatly as vector forces, the magnetic force among, and the resultant motion calculated via vector addition and the rule that the acceleration of the metallic object equals the resultant force divided by the object's mass. But it is a huge leap of faith to suppose that the dust and spider webs between the floorboards can be regimented into this neat picture. The best Cartwright would be prepared to bet is true is that the magnet *could well* lift the earring. And this remains a weak prediction!

³³ Indeed, in most cases it is just because we know both a rule of composition and the contribution of a full *set* of causes towards the effect that we can make sense of the idea of a *contribution* from any one of them.

³⁴ Cf. Anna Alexandrova (2006) "Connecting Economic Models to the Real World", *Philosophy of the Social Sciences* 36: 173-192 and Julian Reiss (2007) *Error in Economics: The Methodology of Evidence-Based Economics*, London: Routledge.

VII. Two Actions, Two Shores, Two Problems

We have distinguished between

1. two actions: hunting causes and using causes taking place on
2. two shores: methodology and policy that cannot be bridged because of
3. two problems: unstable enablers and external validity.

Together these paint a sad picture for causal laws. On current accounts it is either easy to explain why some one or another of our best methods for hunting causal laws should be reliable or it is easy to account for why causal laws can be used for predictions when we propose to change the world. But no account on offer does both at once very well. There are two obvious conclusions:

Conclusion 1: The state of philosophy reflects the state of nature. Causal laws are not worth the paper they are written on. They can be found alright – and it is clear why our best methods for finding them work so well: they just are ‘that which results from this method’. Or they can be used just as we want. But our elaborate methods for testing are neither necessary nor sufficient for claims that give true conclusions about policy manipulations. So we might as well chuck all those elaborate tests.

Conclusion 2: There is a lot of work left for philosophy to do: to find good, rich theories of causality that support method and use in one fell swoop.

We can hope for Conclusion 2, but as always the proof of the pudding will be when we have the theory on our plate and have cut it open to find inside not just one but the two sixpences of hunting and use together.