PROBABILISTIC CAUSALITY AND SIMPSON'S PARADOX*

RICHARD OTTE†

Department of Philosophy
University of California at Santa Cruz

This paper discusses Simpson's paradox and the problem of positive relevance in probabilistic causality. It is argued that Cartwright's solution to Simpson's paradox fails because it ignores one crucial form of the paradox. After clarifying different forms of the paradox, it is shown that any adequate solution to the paradox must allow a cause to be both a negative cause and a positive cause of the same effect. A solution is then given that can handle the form of the paradox that Cartwright's solution ignored, and allows causes to be both a positive and a negative cause of an effect.

Nancy Cartwright (1979) has written a recent article in which she discusses the basic idea of probabilistic causality, namely, that a cause raises the probability of its effect. Cartwright questions this assumption by noting that it is often false that a cause raises the probability of its effect. Cartwright asks us to consider the following example. Let us suppose that smoking causes heart disease, and that exercise prevents heart disease. Since we believe that smoking causes heart disease, we would expect smoking to raise the probability of heart disease, $P(H/S) > P(H)$. However, this may not be true if smoking and exercise are highly correlated, and if exercise is more effective at preventing heart disease than smoking is at causing it. This might be the case if people who like to smoke also like to exercise, and that the result is that they have a lesser chance of having heart disease than a normal person. In this situation we find that $P(H/S) < P(H)$, because most of the smokers are also people who exercise, and exercise is better at preventing heart disease than smoking is at causing it; although smoking causes heart disease, smoking actually lowers the probability of heart disease.

A situation such as this will arise whenever the cause is sufficiently correlated with a third preventative factor of sufficient strength. This is known as *Simpson's paradox*. Simpson's paradox is that any correlation in a population can be reversed in the subpopulations by finding a third

*Received November 1983; revised February 1984.

†The author would like to thank Paul Humphreys, Tom Grimes, and two anonymous referees for this journal for helpful comments on an earlier draft of this paper. This research was supported in part by Faculty Research funds and funds granted by the Regents of the University of California.

Philosophy of Science, 52 (1985) pp. 110–125.

variable that is properly correlated with the other variables. Thus positive relevance can become negative relevance or independence, negative relevance can become positive relevance or independence, and independence can become positive or negative relevance. Almost anything is possible. But if this is true, it seems hopeless to require that causes raise the probability of their effects. In the rest of this paper I will show that Cartwright's treatment of Simpson's paradox is inadequate. I will show that one major reason it is inadequate is because it ignores one of the various forms of the paradox. I then provide a solution which will account for versions of the paradox that Cartwright's solution ignores.

1. Cartwright's Treatment of Simpson's Paradox. In response to Simpson's paradox, Cartwright notices that all of the counterexamples to the claim that causes are positively relevant to their effects depend upon a correlation between the cause and some other causally relevant factor in such a way that the positive relevance is changed. The natural answer to these counterexamples, which Cartwright sees, is to claim that causes increase the probability of their effects when these correlations are absent. The problem is to try to characterize this more explicitly.

In the example we discussed concerning smoking, exercise, and heart disease, we saw that smoking lowered the probability of contracting heart disease. This was because most smokers were also exercisers. Exercising is the causally relevant factor which reverses the positive relevance of smoking to heart disease. However, Cartwright notices that in situations in which there are none of these correlations, smoking always raises the probability of heart disease. If we had a population in which everyone exercised, and thus smoking was not correlated with exercise, we would find that smoking increased the probability of heart disease. Similarly, if we took a population in which nobody exercised, smoking would not be correlated with exercise; in this population we would also find that smoking increased the probability of heart disease. Cartwright believes that in populations which have no correlations between the cause and other causally relevant factors, a cause will always raise the probability of its effect. These populations which have no correlation between the cause and other causally relevant factors are called *causally homogeneous*. Cartwright (1979, p. 423) summarizes this principle by claiming that "C causes E if and only if C increases the probability of E in every situation which is otherwise causally homogeneous with respect to E." This principle claims that if we take the reference class to be causally homogeneous, then a cause always raises the probability of its effect. This is an answer to the problem posed by Simpson's paradox.

In formalizing the above principle, Cartwright appeals to the idea of a state description which Carnap introduced in his work in inductive logic.

A state description is a total or complete description of a possible state of the universe describable by the language. A state description will be a conjunction of simple statements, and each simple statement will either claim that a certain individual has a basic property, or deny that a certain individual has that property. For each basic property a state description will assign either it or its complement to each individual. Thus a state description is a maximal description of the properties that each individual could consistently have.

Cartwright adapts this idea of state descriptions to talk about maximal sets of consistent causal factors. We need to find a way to characterize all of the situations which are causally homogeneous with respect to E. Cartwright begins by defining a complete set of causal factors for E to be the set of all C_i such that either C_i causes E, written as $C_i \rightarrow +E$, or C_i prevents E, written as $C_i \rightarrow -E$. We will write $C_i \rightarrow +E$ or $C_i \rightarrow -E$ as $C_i \rightarrow \pm E$. If we take a possible arrangement of factors of this set minus C, then we have a situation that is causally homogeneous with respect to all causal factors except C. These possible arrangements are generated by conjoining either the members of the set or their complements. A state description K_j is defined as $K_j = \Lambda \pm C_i$ over the set $\{C_i\}$ of all alternative causal factors, where i ranges from 1 to n. These state descriptions are maximal conjunctions of simple causally relevant factors. This will give us 2^n state descriptions. Each state description will be causally homogeneous with respect to E; these are the populations in which a cause must always raise the probability of its effect.

Cartwright (1979, p. 423) then uses the above conception of state descriptions to give a general principle which defines probabilistic causality: CC: $C \rightarrow E$ if and only if $P(E/C \mathbin{\&} K_j) > P(E/K_j)$ for all state descriptions K_j over the set $\{C_i\}$, where $\{C_i\}$ satisfies

$$C_i \in \{C_i\} \supset C_i \rightarrow \pm E \tag{1}$$

$$C \notin \{C_i\} \tag{2}$$

$$(D)[D \rightarrow \pm E \supset (D = C \text{ or } D \in \{C_i\})] \tag{3}$$

$$C_i \in \{C_i\} \supset \neg(C \rightarrow C_i). \tag{4}$$

This principle is not an analysis of $C \rightarrow E$, because the causal relation appears on both sides of the equivalence in the above principle. But it is important because it provides the connection between probability relations and causality.

In principle CC, condition (1) tells us that if anything is a member of $\{C_i\}$, then it is either a cause of $+E$ or of $-E$. There are no causally irrelevant events that are members of $\{C_i\}$. Condition (2) tells us that C is not a member of $\{C_i\}$. This condition is necessary, otherwise

$P(E/C \;\&\; K_j) > P(E/K_j)$ would always fail. If the state description K_j contained C, then the two probabilities would be equal. If the state description contained $\neg C$, then $P(E/C \;\&\; K_j) = 0$, and the inequality would again fail. Thus condition (2) is necessary. Condition (3) tells us that if any event is a cause of $+E$ or $-E$, then it is a member of $\{C_i\}$, unless it is the event C. This condition ensures that $\{C_i\}$ will contain all causally relevant factors. Condition (4) is added to require that the state descriptions do not contain events in the causal chain between C and E; if events in the causal chain between C and E were in the state descriptions, they might screen off C from E, in which case the above principle would be false.

2. Sufficient Causes and Necessary Causes. One problem with condition CC that Cartwright discusses deals with alternative sufficient causes. Suppose that C_j is a sufficient cause of E. Then in some state descriptions $P(E/C \;\&\; K_j) = P(E/K_j) = 1$, and principle CC fails. This example does not even depend upon there being just one alternative sufficient cause; several other causes may become sufficient when conjoined together. C_j and C_k may both be probabilistic causes, but together a sufficient cause of E. In this case some state descriptions will contain both C_j and C_k, which will make the inequality fail, because $P(E/C \;\&\; K_j)$ will be equal to $P(E/K_j)$ when K_j contains both C_j and C_k. The easy way out of this problem is to deny that any sufficient causes exist. This solution is unacceptable, because it initially seems plausible that causal chains will be composed of both deterministic and probabilistic causes. It is too stringent to require that all causes be probabilistic, and a theory of probabilistic causality should be compatible with the existence of deterministic causes.

Cartwright recognizes the problem that arises from the existence of sufficient causes, and gives the following solution. Cartwright (1979, p. 428) modifies the beginning of principle CC to read:

$$C \rightarrow E \text{ if and only if } (j)[P(E/C \;\&\; K_j) > P(E/K_j) \text{ or}$$

$$P(E/K_j) = 1 = P(E/C \;\&\; K_j)] \text{ and } (\exists j)[P(E/K_j) \neq 1]$$

The second conjunct is needed to prevent everything from being the cause of a necessary fact. This modification of Cartwright's is sufficient to solve the problems that arise from alternative sufficient conditions.

Although Cartwright has successfully dealt with alternative sufficient causes, she has not dealt with alternative necessary causes. Suppose that C_j is a necessary cause of E; from this it follows that C_j will be part of half of the state descriptions, and $\neg C_j$ will be part of the other state descriptions. But for any state description K_j which contains $\neg C_j$, $P(E/C \;\&\; K_j) = 0 = P(E/K_j)$, and the inequality fails. We can easily modify CC to account for this problem:

$C \rightarrow E$ if and only if $(j)[P(E/C \,\&\, K_j) > P(E/K_j)$ or
$P(E/C \,\&\, K_j) = 0 = P(E/K_j)$ or
$P(E/C \,\&\, K_j) = 1 = P(E/K_j)]$ and $(\exists j)[0 < P(E/K_j) < 1]$.

This modification will account for cases in which a state description contains either a necessary cause of E or a sufficient cause of E. I believe that the previous modifications are all necessary in order to avoid serious difficulties that present themselves even before we discuss whether the intuition that CC tries to capture is correct.

3. Philosophical Problems with Cartwright's Solution. It appears to me that principle CC is too strong in that there are instances in which a cause does not raise the probability of its effect in every causally homogeneous situation. As an example, I would like to consider an example that Cartwright presents in support of her theory. Ingesting a poisonous acid is a cause of death, ingesting a poisonous alkali may also be a cause of death, but ingesting both a poisonous alkali and acid may not be harmful at all. From this example, Cartwright (1979, p. 428) states three "causal truths":

(1) ingesting an acid without ingesting an alkali causes death;
(2) ingesting an alkali without ingesting an acid causes death; and
(3) ingesting both an alkali and an acid does not cause harm.

Cartwright claims that these statements are consistent with CC. However, a question arises as to whether CC is consistent with the original claim that ingesting a poisonous acid causes death, and that ingesting a poisonous alkali causes death.

In order for the ingestion of acid to be a cause of death, it must raise the probability of death in every causally homogeneous population. But some of these causally homogeneous state descriptions will contain the ingestion of alkali. In these state descriptions, the ingestion of acid will actually lower the probability of death instead of raising it as it should. Thus according to principle CC, the ingestion of poisonous acid is not a cause of death. Similar reasoning will show that according to principle CC, ingestion of poisonous alkali is not a cause of death; both of these results are unintuitive.

I suspect Cartwright would reply to this objection by claiming that the ingestion of acid poison is not a cause of death, but that ingestion of acid poison without ingestion of alkali is a cause of death. Similarly, ingestion of alkali is not a cause of death, but ingestion of alkali without ingestion of acid is a cause of death. I find this very unintuitive. In order to save the claim that a cause raise the probability of its effect in every causally homogeneous state description Cartwright has resorted to requiring that something similar to a total cause be specified. There will be many causal defeaters that can prevent a cause from raising the probability of its effect. I find it unacceptable to require that all of these *possible causal defeaters* be included in the cause. Cartwright herself introduced this example (1979, p. 428) by claiming that "ingesting an acid poison may cause death; so too may the ingestion of an alkali poison." It seems to me that there are cases in which the ingestion of acid does cause death, even though there are situations in which it does not raise the probability of death.

Let us consider another example which illustrates the fact that principle CC requires that all possible causal defeaters be specified. Suppose that we have a laser connected by wires to a power supply. Turning on the power supply does seem to be a cause of the laser firing; this appears to be uncontroversial. However, let us see if turning on the power supply raises the probability of the laser firing in every causally homogeneous state description. One member of the set $\{C_i\}$ will be that there is a break in the wires connecting the laser to the power supply, or that before the power supply is turned on someone cut the wires. These events are causal factors that prevent the laser from firing, and thus will be members of $\{C_i\}$. But if the state description contains the information that the wires are broken, then turning on the power supply does not raise the probability of the laser firing; the power supply has no effect on the laser in this situation. It is important to notice that the probability of the laser firing in this situation is not zero; although it is extremely unlikely, it is physically possible that the laser could fire without power from the power supply. If the probability of the laser firing given that the wires were cut was zero, then principle CC would be able to handle this case. However, in this case the probability of the laser firing given that the wires are cut is not zero; thus principle CC is faulty. In this situation, principle CC would tell us that the cause of the laser firing is turning on the power supply when nobody has cut the wires, and the wires have not been broken, etc. This is a cause that excludes all possible defeaters. But this is an unintuitive and unsatisfactory result. Turning on the power supply is a cause of the laser firing; we do not have to specify a total cause.

A situation like this will arise whenever there is a causal defeater that can render the cause ineffective, and yet the probability of the effect is not zero or one. The above example uses a laser to get an irreducible probability. It could have been done in many other ways. The important point is that a cause can be defeated in such a way that it does not raise the probability of the effect in every causally homogeneous situation. This shows that Cartwright's basic thesis, as expressed in principle CC, is too narrow because it excludes genuine cases of causation.

4. Skyrms's Suggestion. Brian Skyrms discusses the problem of Simpson's paradox, and presents a principle which is a weakened version of principle CC. Skyrms never endorses this principle, but only claims that it is an interesting weakening of Cartwright's proposal. Since we have seen that Cartwright's principle CC was too strong, perhaps Skyrms's weakening of it will be adequate. Skyrms's (1980, p. 108) definition is as follows:

Pareto-Dominance Condition:

$$Pr(E \text{ given } C \,\&\, B_i) \geq Pr(E \text{ given } \neg C \,\&\, B_i) \text{ for every } B_i$$

$$Pr(E \text{ given } C \,\&\, B_i) > Pr(E \text{ given } \neg C \,\&\, B_i)$$
for some B_i in the fundamental partition.

In this condition we can consider the B_i's to be the same as the K_j's in Cartwright's definition. The difference between CC and the Pareto-dominance condition is that whereas CC requires that the cause raise the probability of its effect in every causally homogeneous state description, the Pareto-dominance condition merely requires that a cause not decrease the probability of its effect in any state description, and that it increase the probability in at least one of the state descriptions.

It is interesting to note that Cartwright considered weakening her analysis in such a way that would be similar to the Pareto-dominance condition, but she found that weakening unsatisfactory. In discussing the reasoning why she required that the cause raise the probability of the effect in every state description, and not just decrease the probability of the effect in any of the state descriptions, she (1979, p. 428) says:

> Must a cause increase the probability of its effect in *every* causally fixed situation? Mightn't it do so in some, but not in all? I think not. Whenever a cause fails to increase the probability of its effect, there must be a reason. Two kinds of reasons seem possible. The first is that the cause may be correlated with other causal factors. This kind of reason is taken account of. The second is that interaction may occur. Two causal factors are interactive if in combination they act like a single causal factor whose effects are different from at least one of the two acting separately. For example, ingesting an acid poison. . . .

From the above passage we see that Cartwright rejects the Pareto-dominance condition because she believes that it is not strict enough.

The Pareto-dominance condition can handle some cases of causation that Cartwright's principle CC could not. If we take the example of the laser and the cut wires, which principle CC could not handle, we find that it presents no problem for the Pareto-dominance condition. Principle CC had problems with this example because in state descriptions in which the wires were cut, turning on the power source did not raise the probability of the laser firing. Thus, turning on the power is not a cause of the laser firing, according to principle CC. However, the Pareto-dominance condition does not require that the probability of the laser firing increase in every state description; it only requires that it not decrease in any of the state descriptions and that it increase in at least one state description. The turning on of the power supply does not decrease the probability of the laser firing in any state description, and it certainly raises it in some of them. Thus turning on the power supply is a cause of the laser firing according to the Pareto-dominance condition, which is as it should be.

5. Critical Discussion of the Pareto-Dominance Condition. Problems for the Pareto-dominance condition arise when we realize that there are some state descriptions in which certain events will defeat the cause in such a way that it will actually lower the probability of the effect occurring. Suppose that the power supply was connected to another alternative power supply in such a manner that if the first one is turned on, then the second cannot be turned on, even if the wires between the first power supply and the laser are cut. In this situation the turning on of the first power supply would lower the probability that the laser will fire, which would exclude it from being a cause according to the Pareto-dominance condition.

A similar situation arises in the example that I discussed concerning ingesting acid and ingesting alkali. In that example, there are some state descriptions in which ingesting acid actually lowers the probability of death; namely, those which include the ingestion of alkali. Thus the ingestion of acid is not a cause of death, according to the Pareto-dominance condition. The preceding two examples show that like Cartwright's principle CC, the Pareto-dominance condition is also too strong.

A reply to these objections is available. One could reply that one is not interested in general laws, like Cartwright was, but rather one is interested in actual causal chains. Skyrms (1980, p. 109) expresses this application of the Pareto-dominance condition in the following:

> How does this analysis apply to causal factors for events? Here we want to fix the relevant factors that are in fact present. A heart attack did cause poor Cecil's death. It is true that being run over by a steamroller screens off a heart attack from death, but Cecil was not, in fact, run over by a steamroller. We can neglect those cells which include being run over by a steamroller, and indeed the coroner would like to zero in on that cell which includes the true constellation of background causal factors.

This statement of Skyrms's is an answer to the above objections. Unlike Cartwright, Skyrms is interested in actual causal chains, and not general causal laws. Because of this, he does not need to consider every possible causally relevant factor in the state descriptions. In contrast to this, Cartwright was interested in general causal laws and not in specific causal chains. She is attempting to explicate the relationship between probability and causality in general causal laws such as smoking causes heart disease. Because of this, Cartwright had to require that the state descriptions be formed from a set $\{C_i\}$, which was a set that contained *all* causally relevant factors. With certain exceptions, any factor which can either cause the effect or prevent the effect from occurring, regardless of whether that factor was present in a certain causal chain, was to be included in $\{C_i\}$. The reason that Cartwright could not limit the set $\{C_i\}$ to the factors present, is that she was not dealing with actual chains that were made up of particular events. Thus by dealing with particular causal chains instead of general laws, we can avoid the necessity of including all causally relevant factors in $\{C_i\}$, whether they are present or not.

By ignoring the state descriptions which contain events that are not actual the above problems can be avoided. In the laser example, it is not true that the wires are cut and that there is an alternative power source that is connected in such a way that it cannot be turned on if the other power source is turned on. Thus it would not be a member of the B_i that is conditionalized on. Similarly, if the person did not ingest an alkali along with the acid, then ingesting an alkali would not be a member of the B_i, and thus the ingestion of acid would not lower the probability of death in this situation. The above counterexamples can be handled by dealing with actual causal chains and not with general causal laws.

But if we, as Skyrms suggests, ignore all of the cells or state descriptions that contain causal factors that are not in fact present, then we will always be left with only B_i to conditionalize on. Skyrms's suggestion then becomes just to see if the cause raises the probability of the effect in the reference class that contains all of the relevant causal factors, when we hold these factors fixed. Although Skyrms does not discuss this problem, we must also require that B_i be subject to the same conditions that Cartwright's state descriptions were; otherwise serious problems would result. We shall see that the main problem facing this proposal concerns what factors shall be included in B_i.

6. Three Versions of the Paradox. In order to see exactly why principle CC and the Pareto-dominance condition cannot handle Simpson's paradox, we must distinguish carefully different forms of the paradox. I think that Cartwright is correct when she claims that the reason that Simpson's paradox arises is because a third variable is correlated with the cause and effect in such a way that the relevance relations can change. Looking at this more closely, we can see that there are three basic ways in which a third variable can be correlated with a cause and effect and change the relevance relations; these are diagrammed in figures 1, 2, and 3.[1] In those figures, *C* will stand for the cause, *E* will stand for the effect, and *V* will stand for the third variable that is correlated with them in such a way to reverse the relevance relations. A "+" or a "−" is placed by the causal connections to indicate whether the causal relation is one of positive or negative causation.

One way *V* could be correlated with *C* and *E* would be for *V* to be a probabilistic cause of *C*, and a probabilistic negative cause of *E*. In this

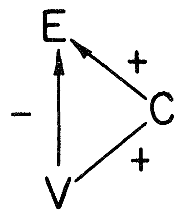

Figure 1

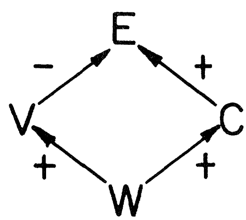

Figure 2

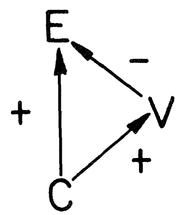

Figure 3

[1]Figures 1, 2, and 3 do not exhaust the possible statistical and causal relations between *C*, *E*, and *V*. There are other cases where *C* is a negative cause of *E*, but because of a correlation between *C* and *V* we find that *C* raises the probability of *E*. Figures 1, 2, and 3 merely give three basic ways that *C* can be a positive cause of *E*, and yet lower the probability of *E* because of a correlation with *V*. I believe that my final analysis of $C \rightarrow E$ will account for those other cases also.

case, which is illustrated by figure 1, the following probability relations might hold:

1. $P(E/C) < P(E)$
2. $P(E/C \& V) > P(E/\neg C \& V)$.

C would lower the probability of E because V is a more effective negative cause of E than C is a positive cause of E. An example of a situation that has this structure is given to us by Skyrms:

> Suppose that air pollution got so bad that most people in the cities refrained from smoking out of sheer terror of putting their lungs in double jeopardy, while many people in areas of the countryside with relatively little pollution felt that they could allow themselves the luxury of smoking. (p. 106)

In this situation living in the countryside, V, is a negative cause of lung cancer, E, and a positive cause of smoking, C. In this situation, smoking will actually lower the probability of getting lung cancer, because smoking is correlated with living in the countryside, which is negatively correlated with lung cancer. This example illustrates a form of Simpson's paradox diagrammed by figure 1.

Another way that V could be correlated with C and E in such a way that equations 1 and 2 would hold would be for there to be some other variable, W, which is a common cause of both V and C. This situation is illustrated by figure 2. An example of this could be constructed from the example given at the beginning of this article concerning exercise, smoking, and heart disease. Let us suppose that the reason that people who like to smoke like to exercise is a certain genetic characteristic that is a common cause of both of those desires. The common cause would account for the strong correlation between those desires, without there actually being a direct causal connection between the two of them.

The third way that I will discuss that C and V could be correlated in such a way that equations 1 and 2 are true would be for C to be a cause of V. This is diagrammed in figure 3. A good example of this is given by Hesslow, who was one of the first philosophers to notice the problem that Simpson's paradox raised for theories of probabilistic causality. Hesslow (1976 p. 291) says that the taking of oral contraceptives (C) can cause thrombosis (E). But since pregnancy can also cause thrombosis, not being pregnant (V) helps prevent thrombosis. The taking of oral contraceptives is a cause of not being pregnant, which is a negative cause of thrombosis. In this situation the taking of oral contraceptives can contribute to the occurrence of thrombosis, and it can cause one not to become pregnant, which helps prevent thrombosis. We could easily construct other examples that have this causal structure. All that is required is that a cause C both cause an effect E and cause some other event V which is a negative cause of E. All three of the above situations I have described are versions of Simpson's paradox, because the negative correlation between the cause and the effect arises as a result of a correlation between them and a third factor.

However, the above does not constitute a solution to Simpson's paradox because it is unclear what events will count as a legitimate third variable and be symbolized by V in the above relations. Obviously there must be some restrictions on V, otherwise the above relations would not hold in the cases in which we wanted them to. For example, if V is a necessary and sufficient cause of C, then the above relations will not be true. So the real work in applying this idea towards a solution of Simpson's paradox is to find a more accurate characterization of the reference class in which the cause must raise the probability of the effect. This is what Cartwright was attempting to do when she introduced the notion of a state description. As Cartwright noted, not every proposition or its negation can be a member of the state description, without the idea of a state description becoming useless for the purposes of solving Simpson's paradox.

For our present purposes, we shall concentrate on one requirement that Cartwright places on the state descriptions. This is the requirement that no effect of C be a member of $\{C_i\}$. As was mentioned earlier, the reason for this requirement is that a member of a causal chain between C and E might screen off C from E. If this happened, then C would be irrelevant to the occurrence of E, and the cause would not raise the probability of the effect. If this requirement is dropped, Cartwright's solution to Simpson's paradox falls apart.

The problem with this requirement is that it seems to exclude cases of Simpson's paradox such as those diagrammed by figure 3. A legitimate case of Simpson's paradox is where the correlation between the cause and the third variable arises because the cause is a cause of the third variable. But when we apply principle CC to these cases, we find that since the other variable is an effect of the cause, that neither it nor its negation can appear in the state descriptions. But if the state descriptions cannot contain either the third variable or its negation, then the cause is still correlated with the other variable, and thus the cause still lowers the probability of the effect. Thus we see that by not allowing the third variable to be in the state descriptions, principle CC does not give the correct answer to the type of situation diagrammed by figure 3. Cartwright's solution can handle cases of Simpson's paradox which have the form diagrammed by figures 1 and 2, but it is unable to handle cases which have the form diagrammed by figure 3.

One might reply to my criticism by claiming that in figure 3 C is not a cause of *E*. It might be claimed that it is a negative cause of *E* and not a positive cause of *E*. The reasoning behind this would be that *C* does in fact lower the probability of *E* occurring. The intuition would be that if *C* occurs and everything else remains fixed, then the probability of *E* occurring is lowered, and if *C* does not occur and everything else is fixed, the probability of *E* is higher.

In response to this objection, I would like to claim that it is a mistake to believe that if something is a negative cause, i.e., it prevents the occurrence of an effect, that it cannot also be a positive cause and contribute to the occurrence of the effect. I think that our intuitions are that a cause can both contribute to the occurrence of an event and contribute to the non-occurrence of the same event. A cause can be a positive and a negative cause of the same event. This is what I think is the proper analysis of the situation in figure 3. In this case, we see that *C* is a positive cause of *E*, which is evident because of the direct connection between *C* and *E*. But *C* is also a negative cause of *E*, because *C* contributes to the occurrence of *V*, and *V* is a negative cause of *E*. Thus in figure 3, *C* both contributes to the occurrence of *E*, and contributes to the non-occurrence of *E*. This is possible because *C* results in two different causal chains which are independent of one another. One of these chains tends to prevent *E* from occurring, and one of them contributes to the occurrence of *E*. I think that this is a good reason to claim that *C* is both a positive and a negative cause of *E*. Most treatments of probabilistic causality do not allow for this possibility. According to them, a cause must be either a positive cause or a negative cause; it cannot be both.

Another objection that has been raised against my claim that a cause can be both a positive and a negative probabilistic cause is as follows. If we look at figure 3, we see that *C* is a direct positive cause of *E*, but the negative causal link from *C* to *E*, through *V*, is indirect. Thus it is improper to claim that *C* is both a positive and a negative cause of *E*, because *C* is an indirect negative cause of *E*.

This objection fails because it assumes that if a cause is an indirect negative cause, then it is not a genuine negative cause. But most causes that we speak of are indirect, and being an indirect cause does not prevent them from also being a positive cause or a negative cause. In the example given, the taking of oral contraceptives (*C*) is not a direct positive cause of thrombosis (*E*). There are intermediate events, but that does not imply that the taking of the contraceptives is not a positive cause of thrombosis. One cannot use the existence of intermediate events to avoid saying some causes are both positive and negative causes without conflicting with many of our intuitions about causation.

If we require that a cause be either a positive cause or a negative cause, but not both, our analysis of the thrombosis case will be unintuitive. In this case we want to say both that the taking of oral contraceptives contributes to the possibility of thrombosis, and that it helps prevent thrombosis. Some women may take oral contraceptives in order to prevent thrombosis. Their reason for doing so would be that oral contraceptives is a negative cause of thrombosis. Other women may not take oral contraceptives because they are concerned that they may cause thrombosis. I think that both of these arguments are reasonable, because we recognize that oral contraceptives can be both a positive and a negative cause of thrombosis. If we were to adopt Cartwright's position we would end up denying the dual nature of many causes.

7. A Solution. Although Cartwright's proposal cannot handle cases of Simpson's paradox that have the form of figure 3, we can give an analysis which is able to handle such cases. We will limit ourselves to Markov chains composed of probabilistic causes. Unlike Cartwright's proposal, we will be dealing with specific causal chains and not general causal laws. Let us first define a *simple cause*:

> * C_t is a simple cause of E if and only if $P(E/A_t \,\&\, C_t) > P(E/A_t \,\&\, \neg C_t)$, where A_t is all of the factors, except C, present at time t that are causally relevant to E.

The information A_t contains all of the causally relevant factors which influence the occurrence of E, with the exception of C. If we had liked, we could have defined A_t differently by including all causally relevant factors except for C that occur at or *earlier than* time t. The only difference that this would make is if temporal action at a distance is possible; I will continue to use the original formulation. Principle * is based on Skyrms's idea that it is only the factors that actually occur that are conditionalized on in determining whether a cause raises the probability of the effect.

Once we have defined a simple cause, we can then define $C \rightarrow E$ as follows:

> ** $C \rightarrow E$ if and only if C is a simple cause of E, or there is an event D such that C is a simple cause of D and D is a simple cause of E.

Principle ** is an extension of principle *. It tells us that C can be a cause of E by either being a simple cause of E according to principle *, or by there being some chain of simple causes which link C to E.[2]

It is important to realize that principle ** is not an analysis of causa-

[2]One reason our analysis is limited to Markov chains is to ensure the transitivity of probabilistic causes. See Eells and Sober (1983) for a discussion of this problem.

tion, because it relies on the notion of a causally relevant factor. Like Cartwright's principle CC, it merely attempts to explicate the relationship that holds between causation and positive statistical relevance. There are also other serious problems, such as the distinction between genuine and spurious causes, and the connection between deterministic and probabilistic causation, which must be solved before we can have an adequate analysis of causation.

One interesting problem with probabilistic causation that principle ** brings out is that if Q is negatively relevant to R and R is negatively relevant to S, then Q will be positively relevant to S. In terms of probabilistic causation, if Q is a negative cause of R and R is a negative cause of S, should we say that Q is a positive cause of S? Intuitions on this situation are mixed, and there does not appear to be an obviously correct answer. My principle ** supports the view that if Q is a negative cause of R, and R is a negative cause of S, then Q is a positive cause of S; Q will be a simple cause of S according to principle *. It will also be the case that Q is a simple cause of $\neg R$, and $\neg R$ is a simple cause of S, which means that Q is a positive cause of S according to principle **. My reasons for having principle ** support this view are as follows. If M is a positive cause of R, and R is a negative cause of S, I think we would claim that M is a negative cause of S. This is because M contributes to the occurrence of a negative cause of S. But Q hinders the occurrence of a negative cause of S. Q tends to prevent R from occurring, which tends to prevent S from occurring; thus it contributes to the occurrence of S, which is why I think it is a positive cause of S. Consider the following example. Suppose that Oblomov has a headache (S) and that Zakhar decides to go to the store for some aspirin to give to Oblomov (R). Chichikov wants Oblomov to have a headache, so he places a thorn in Zakhar's shoe (Q), in order to dissuade Zakhar from going to the store for aspirin. In this situation the thorn in Zakhar's shoe (Q) is a negative cause of Oblomov getting some aspirin (R) which is a negative cause of Oblomov having a headache (S). If we were to ask Chichikov why he put the thorn in Zakhar's shoe, he might reply that he wanted to cause Oblomov's headache to continue. Chichikov would be attempting to cause S to occur by doing Q, which is a negative cause of a negative cause of S; one way to contribute to the occurrence of S would be to hinder the occurrence of a negative cause of S. One might object to my reasoning by claiming that Q failed in its attempt to hinder the occurrence of R, and thus it is improper to view this as a situation in which Q is hindering a negative cause of S. However, I think that if we had a case in which F was a negative cause of G, and G was a positive cause of H, we would say that F was a negative cause of H; it is similar reasoning which leads

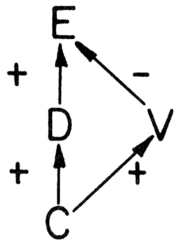

Figure 4

me to think that it is irrelevant whether the negative cause "fails" to produce its effect.[3]

Principle ** can handle all cases of Simpson's paradox that I know of. The first disjunct of ** handles cases that are diagrammed by figures 1 and 2. The second disjunct handles cases diagrammed by figure 3. In a situation diagrammed by figure 3, there will be some event D between C and E such that C raises the probability of D and D raises the probability of E. This is illustrated by figure 4. What principle ** relies upon is V not being causally relevant to the occurrence of D. This is what allows C to be a simple cause of D. We should also notice that principle ** will sometimes yield the result that $C \rightarrow +E$ and $C \rightarrow -E$. I believe that this is the result that we want in cases diagrammed by figure 3, and theories which do not have this result will not do justice to our intuitions concerning such cases. Previous theories have failed to solve Simpson's paradox by failing to account for these situations.

REFERENCES

Cartwright, N. (1979), "Causal Laws and Effective Strategies", *Noûs 13*: 419–37.

Eells, E., and Sober, E. (1983), "Probabilistic Causality and the Question of Transitivity", *Philosophy of Science 50*: 35–57.

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

Humphreys, P. (1980), "Cutting the Causal Chain", *Pacific Philosophical Quarterly 61*: 305–14.

Skyrms, B. (1980), *Causal Necessity*. New Haven: Yale University Press.

[3]For a different analysis of this case see Humphreys (1980).