Probabilistic Measures of Causal Strength[†]

Branden Fitelson

Department of Philosophy

University of California, Berkeley

Christopher Hitchcock

Division of Humanities and Social Sciences

California Institute of Technology

Abstract: A number of theories of causation posit that causes raise the probability of their effects. In this paper, we survey a number of proposals for analyzing causal strength in terms of probabilities. We attempt to characterize just what each one measures, discuss the relationships between the measures, and discuss a number of properties of each measure.

One encounters the notion of 'causal strength' in many contexts. In linear causal models with continuous variables, the regression coefficients (or perhaps the standardized

---

[†] We would like to thank Jim Woodward, the audience at the Second Annual Formal Epistemology Festival (FEF2), and two anonymous referees for useful comments and discussion.

coefficients) are naturally interpreted as causal strengths. In Newtonian Mechanics, the total force acting on a body can be decomposed into component forces due to different sources. Connectionist networks are governed by a system of 'synaptic weights' that are naturally interpreted as causal strengths. And in Lewis's account of 'causation as influence' (Lewis 2000), he claims that the extent to which we regard one event as a cause of another depends upon the degree to which one event 'influences' the other. In this paper, we examine the concept of causal strength as it arises within probabilistic approaches to causation. In particular, we are interested in attempts to measure the causal strength of one binary variable for another in probabilistic terms. Our discussion parallels similar discussions in confirmation theory, in which a number of probabilistic measures of degree of confirmational support have been proposed. Fitelson (1999) and Joyce (MS) are two recent surveys of such measures.

1. Causation as Probability-raising

The idea that causes raise the probabilities of their effects is found in many different approaches to causation. In probabilistic theories of causation, of the sort developed by Reichenbach (1956), Suppes (1970), Cartwright (1979), Skyrms (1980), and Eells (1991), C is a cause of E if C raises the probability of E in fixed background contexts. We form a partition $\{A_1, A_2, A_3 \ldots A_n\}$, where each A_i is a background context. Then C is a cause of E in context A_i just in case $P(E|C \wedge A_i) > P(E|{\sim}C \wedge A_i)$, or equivalently, just in case $P(E|C \wedge A_i) > P(E|A_i)$.[1] The idea is that each background context controls for confounding

[1] Note that both inequalities fail, albeit for different reasons, if $P({\sim}C \wedge A_i) = 0$.

causes of E, so that any correlation that remains between C and E is not spurious. According to Cartwright (1979), each background context should hold fixed (either as being present, or as being absent), every cause of E that is not itself caused by C. Eells (1991) has a similar proposal. If we construct the background contexts in this way, we would expect the conditional probabilities of the form $P(E|C \wedge A_i)$ and $P(E|{\sim}C \wedge A_i)$ to take values of 0 or 1 if E is caused deterministically. However, as Dupré (1984) points out, this carves up the background conditions more finely than is needed if the goal is simply to avoid confounding. For this purpose, it suffices to hold fixed the common causes of C and E. If we construct the more coarse-grained partition in this way, the conditional probabilities $P(E|C \wedge A_i)$ and $P(E|{\sim}C \wedge A_i)$ might take intermediate values even if determinism is true. An issue remains about what it means to say that C causes E *simpliciter*: whether it requires that C raise the probability of E in *all* background contexts (the proposal of Cartwright 1979 and Eells 1991), whether it must raise the probability of E in some contexts and lower it in none (in analogy with Pareto-dominance, the proposal of Skyrms (1980)), or whether C should raise the probability of E in a weighted average of background contexts (this is, essentially, the proposal of Dupré 1984; see Hitchcock 2003 for further discussion). We will avoid this issue by confining our discussion to the case of a single background context.

In his (1986), Lewis offers a probabilistic version of his counterfactual theory of causation. Lewis says that E *causally depends upon* C just in case (i) C and E both occur, (ii) they are suitably distinct from one another, (iii) the probability that E would occur at the time C occurred was x, and (iv) the following counterfactual is true: if C had not occurred, the probability that E would occur would have been substantially less than x.

Lewis takes causal dependence to be sufficient, but not necessary, for causation proper. In cases of preemption or overdetermination, there can be causation without causal dependence. We will largely ignore this complication here. The reliance on counterfactuals is supposed to eliminate any spurious correlation between *C* and *E*. The idea is that we evaluate the counterfactual 'if *C* had not occurred…' by going to the nearest possible world in which *C* does not occur. Such a world will be one where the same background conditions obtain. So common causes of *C* and *E* get held constant on the counterfactual approach, much as they do in probabilistic theories of causation.

The interventionist approach to causation developed by Woodward (2003) can also be naturally extended to account for probabilistic causation. The idea would be that interventions that determine whether or not *C* occurs result in different probabilities for the occurrence of *E*, with interventions that make *C* occur leading to higher probabilities for *E* than interventions that prevent *C* from occurring. The key idea here is that interventions are exogenous, independent causal processes that override the ordinary causes of *C*. Thus even if *C* and *E* normally share a common cause, an intervention that determines whether or not *C* occurs disrupts this normal causal structure and brings *C* or ~*C* about by some independent means.

2. Assumptions

We will remain neutral about the metaphysics of causation, and about the best theoretical approach to adopt. For definiteness, we will work within the mathematical framework of probabilistic theories of causation. Conditional probabilities are simpler and more familiar than probabilities involving counterfactuals or interventions, although the latter

are certainly mathematically tractable (e.g. in the framework of Pearl 2000). We will assume that we are working within one particular background context A_i. Within this context, C and E will be correlated only if C is causally relevant to E. We will leave open the possibility that the context is not specified in sufficient detail to ensure that the conditional probabilities $P(E|C \wedge A_i)$ and $P(E|{\sim}C \wedge A_i)$ take extreme values if determinism is true. To keep the notation simple, however, we will suppress explicit reference to this background context. Moreover, when we are considering more than one cause of E, C_1 and C_2, we will assume that the background condition also fixes any common causes of C_1 and C_2. In addition, we shall assume that C_1 and C_2 are probabilistically independent in this background context. This means that we are ignoring the case where C_1 causes C_2 or vice versa.

In all of our examples, we will assume binary cause and effect variables, X_C and X_E respectively. These can take the values 1 and 0, representing the occurrence or non-occurrence of the corresponding events. We will also write C as shorthand for $X_C = 1$, and ${\sim}C$ as shorthand for $X_C = 0$, and analogously for X_E. We will have a probability function P defined over the algebra generated by X_C and X_E, and also including at a minimum the relevant background context. P represents some type of objective probability. We do not assume that this objective probability is irreducible. For instance, it may be possible to assign probabilities to the outcomes of games of chance, even if the underlying dynamics are deterministic. We leave it open that it may be fruitful to understand causation in such systems probabilistically.

It will often be useful to make reference to a population of individuals, trials, situations, or instances in which C and E are either present or absent. For instance, in a

clinical drug trial, the population is the pool of subjects, and each subject either receives the drug or not. In other kinds of experiments, we may have a series of trials in which C is either introduced or not. Eells (1991, Chapter 1) has a detailed discussion of such populations. We will call the members of such populations 'individuals', even though they may not be people or even objects, but trials, situations, and so on. $P(C)$ is then understood as the probability that C is present for an individual in the population, and likewise for other events in the algebra on which P is defined. This probability is approximated by the frequency of C in the population, although we do not assume that the probability is identical to any actual frequency.

When we discuss counterfactuals, these are to be understood as *non-backtracking* counterfactuals, in the sense of Lewis (1979). The antecedents of these counterfactuals are to be thought of as brought about by small 'miracles' (Lewis 1979) or exogenous interventions (Woodward 2003). We will abbreviate the counterfactual 'if A had occurred, then B would have occurred' by $A > B$. In some cases, we will want to explore the consequences of assuming *counterfactual definiteness*. Counterfactual definiteness is an assumption similar to determinism. It requires that for every individual in a population, either $C > E$ or $C > \sim E$ is true, and either $\sim C > E$ or $\sim C > \sim E$. (This assumption is also called *conditional excluded middle*, and it implies that counterfactuals obey the logic of Stalnaker (1968) rather than Lewis (1973).) If counterfactual definiteness is true, we will assume that holding the relevant background condition fixed

suffices to ensure that $P(E|C) = P(C > E)$ and $P(E|{\sim}C) = P({\sim}C > E)$.[2] We will not, however, assume that counterfactual definiteness is true in general. In particular, counterfactual definiteness seems implausible if determinism does not hold. If counterfactual definiteness is not true, we will assume that holding the relevant background condition fixed ensures that $C > P(E) = p$, where $p = P(E|C)$, and likewise for ${\sim}C$. In other words, if C were the case, then the probability of E would have been p, where p is the actual conditional probability $P(E|C)$.

We are interested in measures of the causal strength of C for E. We will write generically $CS(E, C)$ for this causal strength. Specific measures to be discussed will be denoted by appending subscripts to the function CS. These measures are to be characterized in terms of formulas involving probabilities such as $P(E|C)$, $P(E|{\sim}C)$, and perhaps others as well. It will be convenient to write $CS(E, C)$ to represent the result of applying the mathematical formula to C and E, even if this cannot naturally be interpreted as a causal strength (for example, if C does not raise the probability of E).

When we are considering multiple causes, we will represent the causal strength of C_1 for E in the presence of C_2 as $CS(E, C_1; C_2)$. This will be defined in the same way as CS, but using the conditional probability $P(\bullet|C_2)$ instead of $P(\bullet)$.

We will also be interested in measures of preventative strength, which we will denote $PS(E, C)$. We define the preventative strength of C for E in the following way:

---

[2] Note that we are assuming that C and E do not themselves include counterfactuals. As Lewis (1976) shows, if we allow embeddings, we cannot equate probabilities of conditional with conditional probabilities under pain of triviality.

$$PS(E, C) = -CS(\sim E, C).$$

That is, the preventative strength of C for E is just the causal strength of C for $\sim E$, with a change in sign.[3]

We will consider a variety of candidate measures of causal strength. Some of these have been explicitly proposed as measures of causal strength; others are naturally suggested by various probabilistic approaches to causation. We will discuss the properties of each measure, and try to give an informal explanation of what each one is measuring. Although our overall approach is pluralistic, we will make a few remarks regarding what we take to be the merits and demerits of each measure. We will also discuss the relationships between the measures.

---

[3] This definition assumes that each measure $CS(E,C)$ has a corresponding measure of preventative strength $PS(E,C)$ *with the same functional form* (although replacing E with $\sim E$ in the formula will sometimes result in different terms appearing in the expressions for $CS(E, C)$ and $PS(E, C)$ – see the discussion of continuity properties below). In the recent literature on measures of *confirmational* strength, some authors have proposed that confirmation and disconfirmation should be measured using *different functional forms* (Crupi *et al*, 2007). We will not discuss any such "piece-wise" measures of causal/preventative strength here, but this is an interesting (possible) class of measures that deserves further scrutiny.

For purposes of comparing measures, we will convert all measures to a unit scale. That is, we will adopt the following two scaling conventions for all measures of causal strength (*CS*) and preventative strength (*PS*):

If C causes E, then $CS(E,C) \in (0,1]$.

If C prevents E, then $PS(E,C) = -CS(\sim E, C) \in [-1,0)$.

Measures that are based on *differences* in probabilities will typically already be defined on a [–1,1] scale. But, measures that are based on *ratios* of probabilities will generally need to be re-scaled. We adopt two *desiderata* for any such re-scaling: (a) that it map the original measure onto the interval [–1,1], as described above, and (b) that it yields a measure that is *ordinally equivalent* to the original measure, where $CS_1(E,C)$ and $CS_2(E,C)$ are *ordinally equivalent* iff

For all C, E, C' and E': $CS_1(E, C) \geq CS_1(E', C')$ iff $CS_2(E, C) \geq CS_2(E', C')$.

There are many ways to re-scale a (probabilistic relevance) ratio measure of the form p/q, in accordance with these two re-scaling *desiderata*. Here is a general (parametric) class of such re-scalings, where $\lambda \geq 0$, and $p > q$.[4]

---

[4] We thank Kenny Easwaran for suggesting this general parametric way of representing re-scalings of measures.

$$p/q \rightarrow (p - q) / (p + \lambda\, q)$$

When $\lambda = 0$, we get:

$$p/q \rightarrow (p - q) / p$$

And, when $\lambda = 1$, we have:

$$p/q \rightarrow (p - q) / (p + q)$$

We will discuss several applications of each of these two kinds of re-scalings, below.

3. The Measures

Although we will spend much of the paper introducing the measures in leisurely fashion, we will begin by presenting all of the measures that we will discuss in tabular form. These are shown in table 1. For example, the Eells measure will be represented with a subscript *e*, and defined as the difference in conditional probabilities: $CS_e(E, C) = P(E|C) - P(E|{\sim}C)$.

[Table 1 approximately here]

4. Venn and Boolean Representations

In presenting and discussing the various measures, it will be helpful to represent the probabilities pictorially using Venn diagrams. These will facilitate gaining an intuitive understanding of each measure. Figure 1 represents a situation in which C raises the probability of E. The square has an area of one unit. It represents the entire space of possibilities. This space is divided into six cells. The right side of the rectangle corresponds to the occurrence of C, the left half to $\sim C$. The shaded region corresponds to the occurrence of E. The height of the shaded region on the right hand side corresponds to the conditional probability $P(E|C)$, and the shaded column on the left side corresponds to $P(E|\sim C)$. The two dotted lines are the result of extending the top of each shaded column all the way across the diagram. They are a mathematical convenience: they don't necessarily correspond to any events that are well-defined in the probability space. We will use the lower case letters a through f to denote the six regions in the diagram, and also to represent the areas of the regions. The ratios $a:c:e$ are identical to the ratios $b:d:f$. With this diagram, we can write, for example: $P(C) = b + d + f$; $P(E|\sim C) = e + f$; $P(E|C) - P(E|\sim C) = c + d$; and so on. The representations of the measures in terms of this figure are summarized in table 2.

[Figure 1 approximately here]

[Table 2 approximately here]

Additionally, several of the measures we will discuss can be given simple *Boolean representations.* A Boolean representation for $CS(E, C)$ is a probability space that has the following features:

(a) it includes as events C and E, and two additional events A and Q;

(b) E can be expressed as a Boolean function of the other three events: specifically, $E \equiv A \vee (Q \wedge C)$;

(c) the probabilities on the algebra generated by C and E are the same as the objective probabilities figuring in the measures of causal strength;

(d) $CS(E, C)$ is the (conditional or unconditional) probability of some event in the space involving Q.

Condition (b) is reminiscent of Mackie's definition of an *INUS* condition (Mackie 1974). C is an *INUS* condition for E just in case it is an *insufficient* but *non-redundant* part of an *unnecessary* but *sufficient* condition for E. In the expression $E \equiv A \vee (Q \wedge C)$, C is insufficient for E, since Q must also be present. $Q \wedge C$ is a sufficient condition for E, and C is not redundant: Q alone is insufficient. C is not necessary for E, since A may produce E even in the absence of C. Roughly, we may think of A as the proposition that conditions are right for E to occur in the absence of C, and we may think of Q as the proposition that conditions are right for C to cause E. If determinism is true, we may think of A as representing other causes that are sufficient for E, and of Q as representing the other background conditions that are necessary for C to be a cause of E. However, if there is genuine indeterminism, A and Q will not correspond to any physically real events, but are rather just mathematical conveniences; they may be thought of metaphorically as the results of God's dice rolls. The disjunctive form of the

representation for E in (b), together with its probabilistic nature, has given it the name of a "noisy or" representation.

We will give Boolean representations for four of our measures. These representations differ along two dimensions. First, they differ in the assumptions they make about the probabilistic relations that the new events A and Q bear to C and E and to each other. Second, they identify causal strength with the probabilities of different events, or with probabilities conditional upon different events. The Boolean representations are often helpful for giving an intuitive feel for just what the measures are measuring.

5. The Eells Measure

Eells (1991) offers a probabilistic theory of causation according to which C is a (positive) cause of E just in case $P(E|C \wedge A_i) > P(E|{\sim}C \wedge A_i)$ for every background context A_i.[5] He then defined the 'average degree of causal significance' of C for E as: $ADCS(E, C) = \Sigma_i\, [P(E|C \wedge A_i) - P(E|{\sim}C \wedge A_i)]P(A_i)$.[6] This naturally suggests that when we confine ourselves to a single background context, we define causal strength as:

---

[5] In Eells' theory, causal claims are relativized to a population and a population type. We ignore this complication here.

[6] The proposal of Dupré (1984) that we should count C as a cause of E if it raises the probability of E in a 'fair sample' amounts to the claim that C is a cause of E just in case $ADCS(E, C) > 0$. Interestingly, Eells seems not to have understood this proposal. He was adamantly opposed to Dupré's suggestion and even suggests that it is conceptually confused. In particular, he seems to interpret Dupré's call for averaging over background

$$CS_e(E, C) = P(E|C) - P(E|{\sim}C)$$

This is equal to the area $c + d$ in figure 1. Equivalently, it is the difference between the heights of the two shaded columns. The Eells measure is identical to what psychologists call the *probability contrast* -- *PC* or ΔP for short (see e.g., Cheng and Novick 1990).

The Eells measure may be given a simple Boolean representation. We make the following assumptions about the new events A and Q:

(i) A and Q are mutually exclusive;

(ii) A and C are probabilistically independent; and

(iii) Q and C are probabilistically independent.

As is standard, we identify E with $A \vee (Q \wedge C)$. These assumptions are all shown diagrammatically in figure 2. Given these assumptions, we have:

$$CS_e (E, C) = P(Q).$$ [7]

---

contexts – which is clearly done in the formula for *ADCS* – as equivalent to saying that C causes E just in case $P(E|C) > P(E|{\sim}C)$, where we do not control for confounding factors.

[7] All of the mathematical claims that appear in this paper are verified in a companion *Mathematica* notebook, which can be downloaded from the following URL: http://fitelson.org/pmcs.nb [a PDF version of this notebook is available at

[Figure 2 approximately here]

Intuitively, the Eells measure measures the difference that *C*'s presence makes to the probability of *E*. If we had a population of individuals who all belonged to the relevant background context, and conducted a controlled experiment in which *C* is present for some individuals, and absent in others, the Eells measure would be an estimate of the difference between the relative frequencies of *E* in the two groups.

The Eells measure is related to a concept that statisticians call *causal effect*. Assume counterfactual definiteness, and let X and Y be two quantitative variables. Let x and x' be two possible values of X, and let i be an individual in the population. The causal effect of $X = x$ vs. $X = x'$ on Y for i (abbreviated $CE(Y, X = x, X = x', i)$) is the difference between the value Y would take if X were x and the value Y would take if X were x' for individual i. That is, $CE(Y, X = x, X = x', i) = y - y'$, where $X = x > Y = y$ and $X = x' > Y = y'$ are both true for i. Intuitively, the causal effect is the difference that a hypothetical change from $X = x'$ to $X = x$ would make for the value of Y. Assuming counterfactual definiteness, the Eells measure is the expectation of the causal effect of C vs. $\sim C$ on X_E: $CS_e(E, C) = E[CE(X_E, C, \sim C)]$. For example, if an individual i is such that $C > E$ and $\sim C > \sim E$, then for that individual, the causal effect of C vs $\sim C$ on E is 1. The Eells measure

---

http://fitelson.org/pmcs.nb.pdf]. The companion *Mathematica* notebook makes use of the `PrSAT` *Mathematica* package (Fitelson 2008), which can be downloaded from the following URL: http://fitelson.org/PrSAT/.

corresponds to the expectation of this quantity. On the other hand, suppose that counterfactual definiteness is false. Then the Eells measure is equal to the causal effect of C vs $\sim C$ on the probability of E, or equivalently, the expectation of X_E. Note that while the Eells measure itself is indifferent as to whether counterfactual definiteness is true or false, its interpretation in terms of causal effect is different in the two cases.

The Eells measure is also closely related to what Pearl (2000) calls the *probability of necessity and sufficiency* or *PNS*. Pearl assumes counterfactual definiteness, and defines $PNS(E, C) = P(C > E \wedge \sim C > \sim E)$. Intuitively, $PNS(E, C)$ is the probability that C is both necessary and sufficient for E, where necessity and sufficiency are understood counterfactually. *Monotonicity* is the assumption that $P(C > \sim E \wedge \sim C > E) = 0$. Intuitively, this means that there are no individuals that would have E if they lacked C, and also would have $\sim E$ if they had C. Under the assumption of monotonicity, $CS_e(E, C) = PNS(E, C)$. This is most easily seen by referring to figure 1. Monotonicity is the assumption that no individuals in cell e are such that if they had C, they would be in cell b; and no individuals in cell b are such that if they lacked C, they would be in cell e. Then we can interpret the figure in the following way: e and f comprise the individuals for which $C > E$ and $\sim C > E$; a and b comprise the individuals for which $C > \sim E$ and $\sim C > \sim E$; and c and d comprise the individuals for which $C > E$ and $\sim C > \sim E$. The Eells measure is then the probability that an individual is in the last group. In other words, it is the proportion of the population for which C would make the difference between E and $\sim E$. We reiterate, however, that this interpretation assumes both counterfactual definiteness and monotonicity. In particular, if counterfactual definiteness fails, the Eells measure can continue to take positive values, while *PNS* is identically zero.

The Eells measure exhibits what we might call 'floor effects'.[8] If the background context A_i is one in which E is likely to occur even without C, then this will limit the size of $CS_e(E, C)$: there is only so much difference that C can make. In our Boolean representation, this is reflected in the assumption that A and Q are exclusive. If A is large, then Q must be small. This seems appropriate if we think of causal strength in terms of capacity to make a difference. On the other hand, if we think that the causal strength of C for E should be thought of as the intrinsic power of C to produce E, then it might seem strange that the causal strength should be limited by how prevalent E is in the absence of C.

6. The Suppes Measure

Suppes (1970) required that for C to cause E, $P(E|C) > P(E)$. As we noted above, this is equivalent to the inequality $P(E|C) > P(E|{\sim}C)$. However, the two inequalities suggest different measures of causal strength. Thus we define the Suppes measure as

$$CS_s(E, C) = P(E|C) > P(E)$$

This quantity is equal to the area of region c in figure 1.

---

[8] This terminology is slightly non-standard, since we are describing an upper bound on CS_e rather than a lower bound. However, looking at figure 1, the bound results not from a ceiling that is low, but rather from a floor that is high.

The Suppes measure can be given a simple Boolean representation. Under the same assumptions as those made for the Eells measure, shown in figure 2, we have

$$CS_s\,(E, C) = P(Q \,\&\, \sim C).$$

The Suppes measure is related to the Eells measure as follows:

$$CS_s(E, C) = P(\sim C)CS_e(E, C)$$

[Table 3 approximately here]

Table 3 provides a summary of all the mathematical inter-definitions. Note that we will only explicitly give the expression of a measure in terms of measures that have been previously introduced. The expression of the Suppes measure in terms of e.g. the Galton measure can be derived simply by taking the appropriate inverse: e.g. $CS_s(E, C) = CS_g(E, C)/4P(C)$.

The Suppes measure may be understood operationally in the following way: it is the amount by which the frequency of *E* would increase if *C* were present for all individuals in the population. Indeed Giere (1979) offers a probabilistic theory of causation in which causation is defined in just this way. This way of understanding the Suppes measure is only correct, however, if there is no frequency-dependent causation or inter-unit causation. In biology, mimicry is an example of frequency-dependent causation. For example, the tasty viceroy butterfly protects itself by mimicking the color

patterns of the unpalatable monarch butterfly. But the more prevalent the viceroys become, the less effective this ruse will become. So it may be that among butterflies, mimicking the monarch does in fact raise the probability of survival, but if all butterflies did it, the rate of survival would not go up. For an example of inter-unit causation, consider the effects of second-hand smoke. If everyone were to smoke, lung cancer rates would go up, in part because there would be more smokers, but also because at least some people would be exposed to greater amounts of second-hand smoke. In this case, the Suppes measure would underestimate the amount by which lung cancer would increase. Intuitively, what is going on in each of these cases is that the Suppes measure predicts the amount by which the prevalence of E will change *within a fixed background context*. However, when we increase the prevalence of C in the population, we also change the background context to which at least some members of the population belong. This will have an impact on the prevalence of E that goes beyond that predicted by the Suppes measure within a fixed background context.

The Suppes measure will exhibit floor effects in much the same way the Eells measure does. The Suppes measure is also sensitive to the unconditional value of $P(C)$: for fixed values of $P(E|C)$ and $P(E|{\sim}C)$, $CS_s(E, C)$ decreases as $P(C)$ increases. The feature seems *prima facie* undesirable if we construe causal strength as a measure of the intrinsic tendency or capacity of C to cause E. Such an intrinsic capacity should be independent of the prevalence of C.

7. The Galton Measure

We name this measure after Francis Galton. With quantitative variables X and Y, we often evaluate the relationship between them in terms of the *covariance* or *correlation*. The covariance of two variables is defined as follows:

$$Cov(X, Y) = E(XY) - E(X)E(Y).$$

When X and Y are replaced by the indicator functions X_C and X_E, a little calculation gives us

$$Cov(X_E, X_C) = P(C)P(\sim C)[P(E|C) - P(E|\sim C)].$$

The multiplier $P(C)P(\sim C)$ takes a maximum value of ¼ when $P(C) = .5$, so if we want to convert this measure to a unit scale we will need to normalize. One way to do this is to divide by the standard deviations of X_C and X_E, yielding the *correlation*. We will adopt the simpler expedient of multiplying by 4. Thus:

$$CS_g(E, C) = 4P(C)P(\sim C)[P(E|C) - P(E|\sim C)].$$

This is equal to 4 times the product of c and d in figure 1. The Galton measure is related the Eells and Suppes measures as follows:

$$\begin{aligned} CS_g(E, C) &= 4P(C)P(\sim C)CS_e(E, C) \\ &= 4P(C)CS_s(E, C). \end{aligned}$$

Like the Suppes measure, the Galton measure will exhibit floor effects, and it will be sensitive to the unconditional probability of *C*. The Galton measure intuitively measures the degree to which there is variation in whether or not *E* occurs that is due to variation in whether or not *C* occurs. $CS_g(E, C)$ will take its maximum value when $P(E|C)$ is close to 1, $P(E|{\sim}C)$ is close to 0, and $P(C)$ is close to .5. In these circumstances, $P(E)$ will be close to .5, so there is a lot of variation in the occurrence of *E* – sometimes it happens, sometimes it doesn't. When *C* occurs, there is very little variation: *E* almost always occurs; and when *C* doesn't occur, *E* almost never occurs. So there is a lot of variation in whether or not *E* occurs precisely because there is variation in whether or not *C* occurs. By contrast, suppose that $P(C)$ is close to 1. Then any variation in whether or not *E* occurs will almost all be due to the fact that $P(E|C)$ is non-extreme: *E* sometimes happens in the presence of *C*, and sometimes it doesn't. Likewise if $P(C)$ is close to 0. For example, it might be natural to say that small pox is lethal: it is a potent cause of death. So we might think that the causal strength of small pox for death is high. But the Galton measure would give it a low rating, perhaps even 0, since none of the actual variation in who lives and who dies during a given period is due to variation in who is exposed to small pox: thankfully, no one is any more.

Note that the standard measure of *heritability* used in genetics and evolutionary biology is essentially a measure of correlation, and behaves much like the Galton measure. Because of the sensitivity of the heritability measure to the absolute level of variation in some trait among the parents in a population, heritability is a poor measure of

the intrinsic tendency of parents to produce offspring that resemble them with respect to the trait in question.

8. The Cheng Measure

The psychologist Patricia Cheng proposed that we have a concept of 'causal power', and that this explains various aspects of our causal reasoning (1997). Under the special assumptions we have made, causal power reduces to the following formula:

$$CS_c(E, C) = [P(E|C) - P(E|\sim C)]/P(\sim E|\sim C)$$

In our pictorial representation (figure 1), this is equal to the ratio $d/(b + d)$.

It is well-known that the Cheng measure has a "noisy or" representation. (See, e.g., Glymour 1998). We make the following assumption:

A, Q, and C are both pairwise and jointly independent.

As always, E is identified with $A \vee (Q \wedge C)$. These assumptions are shown schematically in figure 3. Then we can identify

$$CS_c(E, C) = P(Q)$$

Note that while both CS_e and CS_c are identified with $P(Q)$, the probabilistic assumptions underlying the two representations are different.

[figure 3 approximately here]

The Cheng measure is related to our other measures by the following formulae:

$$CS_c(E, C) = CS_e(E, C)/P(\sim E|\sim C)$$

$$= CS_s(E, C)/P(\sim E \wedge \sim C)$$

$$= CS_g(E, C)/4P(C)P(\sim E \wedge \sim C)$$

Only the first of these is particularly intuitive. One way of thinking about the Cheng measure is that it is like the Eells measure in focusing on the difference $P(E|C) - P(E|\sim C)$, but eliminates floor effects by dividing by $P(\sim E|\sim C)$. The idea is that it is only within the space allowed by $P(\sim E|\sim C)$ that C has to opportunity to make a difference for the occurrence of E, so we should rate C's performance by how well it does within the space allowed it.

Cheng conceives of her causal power measure in the following way. Assume that E will occur just in case C occurs and 'works' to produce E, or some other cause of E is present and 'works' to produce E. In our Boolean representation, shown in figure 3, Q corresponds to *C's* 'working', and A corresponds to some other cause's working. $CS_c(E, C)$ is then the probability that C 'works'. These 'workings' are not mutually exclusive: it is possible that C is present and 'works' to produce E, and that some other cause also 'works' to produce E. Thus Cheng's model is compatible with causal overdetermination. A high probability for E in the absence of C needn't indicate that C isn't working most of

the time when it is present. But this is at best a heuristic for thinking about causal power. The nature of this 'working' is metaphysically mysterious. If the underlying physics is deterministic, then perhaps we can understand *C*'s 'working' as the presence of conditions that render *C* sufficient for *E* (represented by *Q* in our Boolean representation). If the causal relationship is indeterministic, however, it is hard to see what this 'working' could be. *C* and various other causes of *E* are present. In virtue of their presence *E* has a certain probability of occurring. On most conceptions of indeterministic causation, that is all there is to the story. (See, e.g. Lewis 1986 and Humphreys 1989, sections 10 and 11; Woodward (1990) challenges this conception. See also Hitchcock (2004) for discussion of the two different models.)

The Cheng measure is related to what Pearl (2000) calls the *probability of sufficiency* or *POS*. Assuming counterfactual definiteness, Pearl defines $POS(E, C) = P(C > E|{\sim}C \wedge {\sim}E)$. That is, in cases where neither *C* nor *E* occur, $POS(E, C)$ is the probability that *E* would occur if *C* were to occur. Conditioning on ${\sim}C \wedge {\sim}E$ means that we are in the rectangle occupied by *a* and *c* in figure 1. Now assume monotonicity: that no individuals in region *e* would move to *b* if *C* were to occur, and no individuals in *b* would move to *e* if *C* did not occur. Then the result of hypothetically introducing *C* to the individuals in region *a* and *c* is to move them straight over to the right hand side. So the proportion of individuals in regions *a* and *c* that will experience *E* when *C* is introduced is equal to $d/(b + d)$. So under the assumptions of counterfactual definiteness and monotonicity, $CS_c(E, C) = POS(E, C)$. If counterfactual definiteness does not hold, however, this interpretation cannot be employed. In this case, CS_g may still take positive values, while *POS* is identically zero.

The Cheng measure does not exhibit floor effects, and it is not sensitive to the absolute value of $P(C)$. For this reason it is a more plausible measure of the intrinsic capacity of C to produce E than any of the others we have discussed.

9. The Lewis Ratio Measure

In formulating the probabilistic extension of his counterfactual theory of causation, Lewis (1986) required that in order for E to be causally dependent upon C, the probability that E would occur if C had not occurred had to be *substantially less* than the actual probability of E. Lewis then remarks that the size of the decrease is measured by the *ratio* of the quantities, rather than their difference. This naturally suggests the following measure:

$$CS_{lr}(E, C) = P(E|C)/P(E|\sim C).$$

This is the ratio $(d + f)/f$ in figure 1. The Lewis ratio measure is equivalent to the quantity called 'relative risk' in epidemiology and tort law: it is the risk of experiencing E in the presence of C, relative to the risk of E in the absence of C (see Parascandola (1996) for a philosophically sensitive discussion of these topics).

The Lewis ratio measure rates causes on a scale from one to infinity (and it gives numbers between zero and one when $P(E|C) < P(E|\sim C)$). Thus if we want to compare it directly with our other measures we will need to convert it to a unit scale. As discussed above, there are a number of ways of doing this. We will consider two. The first, corresponding to setting $\lambda = 1$ in our parametric re-scaling formula above, is:

$$CS_{lr1}(E, C) = [P(E|C) - P(E|\sim C)]/[P(E|C) + P(E|\sim C)]$$

This is equal to $d/(d + e + f)$ in figure 1. This re-scaling of the Lewis ratio measure is related to the Eells measure as follows:

$$CS_{lr1}(E, C) = CS_e(E, C)/[P(E|C) + P(E|\sim C)]$$

Its mathematical relationship to the other measures is insufficiently elegant to be illuminating.

The second re-scaling corresponds to setting $\lambda = 0$:

$$CS_{lr2}(E, C) = [P(E|C) - P(E|\sim C)]/P(E|C)$$

This is the ratio $d/(d + f)$ in figure 1. This re-scaling of the Lewis measure can be given a Boolean representation, using the same probabilistic assumptions as those used for the Eells and Suppes measures (shown in figure 2). Then we have:

$$CS_{lr2}(E, C) = P(Q|\ C \wedge E)$$

This re-scaling is related to our other measures via the following formulae:

$$CS_{lr2}(E, C) = CS_e(E, C)/P(E|C)$$

$$= CS_s(E, C)/P(E|C)P(\sim C)$$

$= CS_g(E, C)/P(E \wedge C)P(\sim C)$

$= CS_c(E, C)[P(\sim E|\sim C)/P(E|C)]$

$CS_{lr2}(E, C)$ is equivalent to the quantity called the *probability of causation* in epidemiology and tort law. It is also related to what Pearl (2000) calls the *probability of necessity*, or *PN*. It will be helpful to consider the latter connection first. Assuming counterfactual definiteness, Pearl defines $PN(E, C) = P(\sim C > \sim E|C \wedge E)$. That is, given that *C* and *E* both occurred, $PN(E, C)$ is the probability that *C* is necessary for *E*, where necessity is understood counterfactually. If we assume monotonicity, then $PN(E, C) = CS_{lr2}(E, C)$. The idea is if *C* and *E* both occur, we are in the region *d & f* in figure 1. Under the assumption of monotonicity, the effect of hypothetically removing *C* will be to shift individuals straight to the left. Thus the proportion of those in region *d & f* that would no longer experience *E* if *C* did not occur would be $c/(c + e) = d/(d + f)$. If we define causation directly in terms of (definite) counterfactual dependence, as is done in the law, then $CS_{lr2}(E, C)$ is the probability that *C* caused *E*, given that *C* and *E* both occurred: hence the name 'probability of causation.' In our Boolean representation, *Q* can be thought of as *C*'s being necessary for *E*, or *C*'s causing *E*. 'Probability of causation' is important in tort law. In civil liability cases, the standard of evidence is 'more probable than not'. Thus if a plaintiff has been exposed to *C*, and suffers adverse reaction *E*, in order to receive a settlement she must establish that the probability is greater than one-half that *C* caused *E*. This is often interpreted as requiring that the 'probability of causation' is greater than .5.

It is worth remembering, however, that the interpretation of $CS_{lr2}(E, C)$ as the probability that C caused E depends upon three assumptions. The first is that counterfactual dependence is necessary for causation. This assumption fails in cases of preemption and overdetermination. We have chosen to ignore these particular problems, although as we have seen, the Cheng measure seems to be compatible with causal overdetermination. The second assumption is monotonicity. The third, and most important, is counterfactual definiteness. If counterfactual definiteness fails, then all we can say about those individuals that experience both C and E is that if C had not occurred, the probability of E would have been p, where p is $P(E|\sim C)$. Thus it is true for *all* the individuals that experience both C and E that the probability of E would have been lower if C had not occurred. So to the extent that there is a 'probability of causation', that probability is 1: for all the individuals that experience both C and E, C was *a* cause of E (although there may be other causes as well). This is how Lewis himself interprets indeterministic causation (Lewis 1986).[9]

Like the Eells, Suppes, and Galton measures, the Lewis ratio measure and its rescalings will exhibit floor effects. Like the Eells and Cheng measures, the Lewis ratio measures and its re-scalings are not sensitive to the unconditional probability of C.

10. The Good Measure

Good (1961 – 2) sought to define a measure $Q(E, C)$ of the *tendency of C to cause E*. The measure he ultimately proposed was $Q(E, C) = log[P(\sim E|\sim C)/P(\sim E|C)]$. We propose to

---

[9] See also the discussion in Parascandola (1996) and Hitchcock (2004).

simplify this formula (in a way that does not affect its ordinal scale) by not taking the log (or equivalently, raising the base (*e* or 10) to the power of *Q*). Since we have already used the subscript '*g*' for the Galton measure, we will use Good's well-known first initials '*ij*'.

$$CS_{ij}(E, C) = P(\sim E|\sim C)/P(\sim E|C)$$

This is equal to the ratio $(b + d)/d$ in figure 1. The Good measure is related to the Lewis ratio measure via the formula:

$$CS_{ij}(E, C) = CS_{lr}(\sim E, \sim C)$$

Like the Lewis ratio measure, the Good measure yields a scale from one to infinity when $P(E|C) > P(E|\sim C)$, and from zero to one otherwise. So we will consider two re-scalings.

$$CS_{ij1}(E, C) = [P(\sim E|\sim C) - P(\sim E|C)]/[P(\sim E|\sim C) + P(\sim E|C)]$$

This is equal to the ratio $d/(2b + d)$ in figure 1. This re-scaling is related to other measures via the following formulae:

$$CS_{ij1}(E, C) = CS_{lr1}(\sim E, \sim C)$$
$$= CS_e(E, C)/[P(\sim E|C) + P(\sim E|\sim C)]$$

It mathematical relationship to the other measures is insufficiently elegant to be illuminating. The second re-scaling is:

$$CS_{ij2}(E, C) = [P(\sim E|\sim C) - P(\sim E|C)]/P(\sim E|\sim C)$$

Which is equal to $d/(b + d)$. Interestingly, this second re-scaling of the Good measure is identical to Cheng measure. Obviously, then, this re-scaling will have the same properties, and be susceptible to the same interpretations, as the Cheng measure. Since the original Good measure and the first re-scaling are ordinally equivalent to the second re-scaling, they will be ordinally equivalent to the Cheng measure and also share many of its properties. Here are some other equivalences involving the second re-scaling of the Good measure:

$$\begin{aligned} CS_{ij2}(E, C) &= CS_c(E, C) \\ &= CS_e(E, C)/P(\sim E|\sim C) \\ &= CS_s(E, C)/P(\sim E \wedge \sim C) \\ &= CS_g(E, C)/4P(C)P(\sim E \wedge \sim C) \\ &= CS_{lr2}(\sim E, \sim C) \end{aligned}$$

11. Other Measures

It is fairly easy to generate other candidate measures. One would be the difference between the Eells and the Suppes measures, namely:

$$CS(E, C) = P(E) - P(E|{\sim}C)$$

This could be understood operationally as the amount by which the frequency of *E* would decline if *C* were completely eliminated (modulo worries about frequency dependent and inter-unit causation). We might think of this as the extent to which *C* is in fact causing *E*. Noting that the Lewis ratio measure is simply the ratio of the two quantities whose difference is the Eells measure, we could define a measure that is the ratio of the two quantities whose difference is the Suppes measure:

$$CS(E, C) = P(E|C)/P(E)$$

And of course we could then take different re-scalings of this measure to convert it to a unit scale. We could also construct an analog of the Cheng measure that makes use of the difference that figures in the Suppes measure:

$$CS(E, C) = [P(E|C) - P(E)]/P({\sim}E)$$

And so on. Since the measures that we have already discussed are more than enough to keep us busy, we will leave an exploration of the properties of these new measures as an exercise for the reader.[10]

12. Properties and Comparisons

In the remaining sections, we will explore some further properties of the measures that we have introduced, and examine some relationships between them. First, we will consider whether any of our measures are ordinarily equivalent, or partially ordinally equivalent. Second, we will examine a number of continuity properties of measures – these involve the behaviors of the measures as $P(E|C)$ decreases from a value greater than $P(E|\sim C)$ to a value less than $P(E|\sim C)$. Finally, we will examine what the measures tell us about causal independence, and compare the independence judgments of the various measures.

13. Ordinal Relationships Between Measures

Our two re-scalings of the Lewis ratio measure are, by design, ordinally equivalent to the original Lewis ratio measure, and to each other. Likewise for the re-scalings of the Good Measure. Moreover, as we have already seen, one of our rescalings of Good's measure is *numerically identical* to Cheng's measure.

---

[10] The computational tools developed in the companion *Mathematica* notebook (see footnote 7) are quite general, and can be applied to various other possible measures of causal strength, and various other properties of measures as well.

$$CS_{ij2}\ (E,\ C) = CS_c\ (E,\ C)$$

Apart from these cases, no other pair of measures we're discussing here are numerically equivalent. Indeed, it turns out that no other pair of measures we're discussing here are *ordinally* equivalent (in general). But, some other pairs of measures are ordinally equivalent *in special types of cases*. Consider the following two special types of cases:

I. Cases involving a single effect (E) and two causes (C_1 and C_2).

II. Cases involving a single cause (C) and two effects (E_1 and E_2).

If two measures (CS_1 and CS_2) are such that, for all E, C_1 and C_2:

$$CS_1(E,\ C_1) \geq CS_1(E,\ C_2) \text{ iff } CS_2(E,\ C_1) \geq CS_2(E,\ C_2)$$

then CS_1 and CS_2 are ordinally equivalent *in all cases of Type I* (or "I-equivalent", for short). And, if CS_1 and CS_2 are such that, for all C, E_1 and E_2:

$$CS_1(E_1,\ C) \geq CS_1(E_2,\ C) \text{ iff } CS_1(E_1,\ C) \geq CS_1(E_2,\ C)$$

then CS_1 and CS_2 are ordinally equivalent *in all cases of Type II* (or "II-equivalent", for short). Various pairs of measures (which are not ordinally equivalent in general) turn out to be either I-equivalent or II-equivalent. For example, the Eells, Suppes, and Galton

measures are all II-equivalent. This can be seen readily by examing the identities in table 3. For a fixed *C*, the Eells, Suppes, and Galton measures are all fixed multiples of one another. Thus, for a fixed *C*, they will agree on comparative judgments of causal strength. Table 4 summarizes all ordinal relationships between measures (a "G-E" in a cell of Table 4 means that the two measures intersecting on that cell are *generally* ordinally equivalent, a "I-E" means they are I-equivalent, and a "II-E" means they are II-equivalent).

14. Continuity Properties of Measures

Some of our measures exhibit the following *continuity* between causation and prevention ("Causation-Prevention Continuity"):

(CPC) $CS(E,C) = -CS(\sim E,C)$.

Recall that we are *defining* $PS(E,C)$ as $-CS(\sim E,C)$. As such, we can also express (CPC) as asserting that the *absolute value* of $CS(E,C)$ is the same as the *absolute value* of $PS(E,C)$. If a measure satisfies (CPC), then we can plug probabilities into the measure without regard to whether *C* causes *E* or prevents *E*. If the measure yields a positive value, that is the causal strength of *C* for *E*; if it yields a negative value, that is the preventative strength of *C* for *E*. By contrast, if a measure does not satisfy (CPC), then we must first determine whether *C* causes *E* or prevents *E* before we know which probabilities to plug into the formula. If a measure violates (CPC), it would suggest that causation and prevention are somehow conceptually different – there is a 'discontinuity'

where $P(E|C) = P(E|{\sim}C)$.[11] For example, the Eells measure is simply the difference between $P(E|C)$ and $P(E|{\sim}C)$. The effect of switching E and ~E is simply to reverse the sign. We can continue to use the same formula regardless of whether $P(E|C) > P(E|{\sim}C)$ or $P(E|C) < P(E|{\sim}C)$. The Suppes and Galton measures similarly obey (CPC). By contrast, the Cheng measure of the causal strength of C for E includes the term $P({\sim}E|{\sim}C)$ in its denominator. Thus if C prevents E, and we want to assess $PS_c(E, C) = -CS_c({\sim}E, C)$, we will need to replace $P({\sim}E|{\sim}C)$ in the denominator with $P(E|{\sim}C)$, as well as merely changing the sign. So except for the special case where $P({\sim}E|{\sim}C) = .5$, we will need to know whether C causes E or prevents E in order to know how to use the formula correctly.

Some measures exhibit the following continuity between causation and omission ("Causation-Omission Continuity"):

(COC) $CS(E, C) = -CS(E, {\sim}C)$.

$CS(E,{\sim}C)$ may be thought of as the causal strength with which the omission or absence of C causes E. If a measure satisfies (COC), then, when C prevents E, $CS(E, C)$ will give us a measure of the extent to which the absence of C causes E (with the sign reversed).Thus such a measure may be thought to treat causation and causation by omission as on a par. For example, the Eells measure satisfies (COC): swapping ~C for C has the effect of

---

[11] We do not mean a literal discontinuity. All of our measures will take the value 0 when $P(E|C) = P(E|{\sim}C)$, and will approach this value from below and above.

switching the two terms, resulting in a change of sign. The Galton measure also satisfies (COC).

Interestingly, one of our re-scalings of the Lewis ratio measure satisfies (COC) while the other does not; similarly for the Good measure. This suggests that the choice of re-scaling will make a substantive difference to how the measures treat causation by omission. It also suggests that there is more to re-scaling than simply preserving ordinal equivalence.

Finally, some measures exhibit the following continuity between causation, prevention, and omission ("Causation = Prevention by Omission"):

(CPO) $CS(E, C) = CS(\sim E, \sim C)$.

Given our definition of *PS*, (CPO) says that the causal strength of *C* for *E* is equal in magnitude and opposite in sign to the preventative strength of ~*C* for *E*. It is easy to see that (CPO) is a logical consequence of the conjunction of (CPC) and (COC). So, any measure that satisfies both (CPC) and (COC) must also satisfy (CPO). But, the converse does not hold. That is, (CPO) is strictly weaker than (CPC) & (COC).[12] As reported in Table 5, the Eells and Galton measures satisfy both (CPC) and (COC). As a result, they both satisfy (CPO) as well. None of our other measures satisfy (CPO). Table 5

---

[12] See (Eells & Fitelson, 2002) for a discussion of these (and other) formal continuity properties of probabilistic relevance measures (in the context of *confirmation*).

summarizes the behavior of our measures of causal strength, with respect to these three continuity properties (see §5 of Eells & Fitelson 2002 for a formally similar table).

[Table 5 roughly here]

15. Causal Independence

Causes sometimes operate *independently* of one another, and sometimes they do not. In this section, we will introduce a notion of causal independence and discuss some of its properties (vis-à-vis the measures of causal strength we are studying). First, we need a way of characterizing when two causes C_1 and C_2 of an effect E operate independently of one another (regarding E). The intuitive idea behind our formal definition of causal independence is that C_1 and C_2 are *independent in causing E* just in case the causal strength of C_1 for E does not depend on whether or not C_2 is also present, and *vice versa*. This is *not* to say that C_1 and C_2 are (probabilistically) independent *of each other*.[13] Formally, this intuitive idea is best captured by the following definition:

---

[13] It is true that we are assuming (for simplicity) that C_1 and C_2 are probabilistically independent, relative to the background context. But, conceptually, this assumption is distinct from the assumption of the causal independence of C_1 and C_2 vis-à-vis E. A similar distinction needs to be made in the context of *confirmational* independence of two pieces of *evidence*, regarding a *hypothesis*. Various accounts of confirmational independence mistakenly conflate these two notions. See (Fitelson 2001, chapter 3).

C_1 and C_2 are *independent in causing E*, according to a measure of causal strength *CS* iff $CS(E, C_1; C_2) = CS(E, C_1; \sim C_2)$.

We will abbreviate this relation $I_{CS}(E, C_1, C_2)$. To avoid embedded subscripts, we will use I_n to label the independence relation generated by CS_n. Because we are assuming that C_1 and C_2 are probabilistically independent (given the background condition), the following two basic facts can be shown to hold — for *all* of our measures of causal strength *CS* (assuming each of C_1, C_2 causes E):

- $I_{CS}(E, C_1, C_2)$ iff $I_{CS}(E, C_2, C_1)$. [I_{CS} is *symmetric* in C_1, C_2.]
- $I_{CS}(E, C_1, C_2)$ iff $CS(E, C_1; C_2) = CS(E, C_1)$. [$I_{CS}$ can be defined in terms of the *absence* of C_2, or just in terms of conditional *vs* unconditional *CS*-values.]

While all of our measures converge on these two fundamental properties of I_{CS}, there are also some important *divergences* between our *CS*-measures, when it comes to I_{CS}.

First, we will consider whether it is possible for various pairs of distinct *CS*-measures to *agree* on judgments of causal independence. That is, for which pairs of measures CS_1, CS_2 can we have *both* $I_{CS_1}(E, C_1, C_2)$ *and* $I_{CS_2}(E, C_1, C_2)$? It should be apparent that ordinal equivalence is sufficient for agreement in independence judgments, although it is not necessary. It follows that the different re-scalings of the Lewis ratio measure will always agree on their independence judgments, as will the different re-scalings of the Good measure. Moreover, the Good measure and its re-scalings yield all the same independence judgments as the Cheng measure. Interestingly, among all the

measures we're discussing here, not all pairs *can* agree on I_{CS}-judgments (apart from the trivial cases where one of C_1 or C_2 is not a cause of E). And, those pairs of measures that *can* agree on *some* I_{CS}-judgments, *must* agree on *all* I_{CS}-judgments. Table 6 summarizes these I_{CS}-agreement results.

[Table 6 roughly here]

Second, we will consider whether a measure *CS*'s judging that $I_{CS}(E, C_2, C_1)$ places *substantive constraints* on the *individual* causal strengths $CS(E, C_1)$, $CS(E, C_2)$. Interestingly, some measures *CS* are such that $I_{CS}(E, C_2, C_1)$ *does* impose substantive constraints on the values of $CS(E, C_1)$, $CS(E, C_2)$. Specifically, the Eells, Suppes, and Galton measures all have the following property:

(†) If $I_{CS}(E, C_2, C_1)$, then $CS(E, C_1) + CS(E, C_2) \leq 1$.

Moreover, *only* the Eells, Suppes, and Galton measures have property (†). None of the other measures studied here are such that $I_{CS}(E, C_2, C_1)$ places such a substantive constraint on the values of $CS(E, C_1)$, $CS(E, C_2)$ for independent causes. (†) Strikes us as an undesirable property: it seems to indicate that there are *a priori* restrictions on which kinds of causes can act independently of one another.

Finally, we ask whether "the conjunction of two independent causes is better than one". More precisely, we consider the following question: which of our measures satisfy the following property for conjunctions of independent causes:

(S) If $I_{CS}(E, C_2, C_1)$, then $CS(E, C_1 \wedge C_2) > CS(E, C_i)$, for both $i = 1$ and $i = 2$.

The intuition behind (S) is that if C_1 and C_2 are independent causes of E, then their conjunction should be a stronger cause of E than either individual cause C_1 or C_2. It is interesting to note that some of our measures *appear* to violate (S).[14] That is, if we think of (S) in *formal* terms, then measures like Eells and Cheng appear to violate (S). The problem here lies with the proper way to unpack "$CS(E, C_1 \wedge C_2)$" for measures like Eells and Cheng, which compare $P(E|C)$ *and* $P(E|{\sim}C)$. When calculating $CS(E, C_1 \wedge C_2)$ for such measures, we should not simply compare $P(E|\ C_1 \wedge C_2)$ and $P(E|\ {\sim}(C_1 \wedge C_2))$, since that involves *averaging* over different possible *instantiations* of causal factors that might undergird the truth of "${\sim}(C_1 \wedge C_2)$". Rather, we should compare $P(E|\ C_1 \wedge C_2)$ and $P(E|\ {\sim}C_1 \wedge {\sim}C_2)$. Thus, for example, for the Eells measure, we would have $CS_e(E, C_1 \wedge C_2) = P(E|\ C_1 \wedge C_2) - P(E|\ {\sim}C_1 \wedge {\sim}C_2)$ Once we correct for this misleading way of unpacking

---

[14] It is important to note here that *all* probabilistic relevance measures of degree of causal strength *must* satisfy the following, *weaker*, *qualitative* variant of (S):

(S_0) If $I_{CS}(E, C_2, C_1)$, then $CS(E, C_1 \wedge C_2) > 0$ [*i.e.*, $C_1 \wedge C_2$ *is a* cause of E].

And, this will be true on *either* way of unpacking "$CS(E, C_1 \wedge C_2)$" discussed below.

"$CS(E, C_1 \& C_2)$" in (S), then it follows that *almost*[15] *all* of our measures of causal strength satisfy (S).

Note that if we redefine $CS(E, C_1 \wedge C_2)$ in this way, then some of the identities in Table 3 will not hold for conjunctive causes. For instance, the identity $CS_s(E, C) = P(\sim C)CS_e(E, C)$ relating the Eells and the Suppes measure for atomic causes, is not preserved. That is, it will not be the case that either $CS_s(E, C_1 \wedge C_2) = P(\sim(C_1 \wedge C_2))CS_e(E, C_1 \wedge C_2)$ or $CS_s(E, C_1 \wedge C_2) = P(\sim C_1 \wedge \sim C_2)CS_e(E, C_1 \wedge C_2)$ in general. Moreover, the redefinition of $CS(E, C_1 \wedge C_2)$ entails that in order to calculate causal strengths, we must identify the appropriate level of atomic causes. Most of the results in this paper have to do only with such atomic (or fundamental/primitive) causal factors

---

[15] This question is particularly difficult to analyze for the Galton measure. We haven't been able to find any plausible re-definition of $CS_g(E, C_1 \wedge C_2)$ which ensures the satisfaction of (S) for the Galton measure. We suspect that the anomalous result occurs for CS_g because of the way we are trying to force what is essentially a *covariation* measure into a measure designed for binary random variables. Intuitively, from a perspecitive of covariation, it makes more sense to somehow think of '$C_1 \wedge C_2$' as a four-valued random variable. Considered just as a binary variable, it stands to reason that sometimes variation in whether or not '$C_1 \wedge C_2$' occurs won't capture some of the variation in whether E occurs, since some of the latter is due to variation in the different ways $\sim(C_1 \wedge C_2)$ can occur. This is a nice illustration of the subtlety of combining the causal strengths of individual ("atomic") causal factors.

(and that is the intended domain for Table 3). The general problem of *combining* atomic causal factors into complex causal factors is a subtle one, which is beyond the scope of the present discussion.

Finally, we note that with this new definition of $CS(E, C_1 \wedge C_2)$, several of our measures yield fairly simple expressions for $CS(E, C_1 \wedge C_2)$ in terms of $CS(E, C_1)$ and $CS(E, C_2)$ in the case of independence:

$I_e(E, C_1, C_2)$ implies $CS_e(E, C_1 \wedge C_2) = CS_e(E, C_1) + CS_e(E, C_2)$

$I_s(E, C_1, C_2)$ implies $CS_s(E, C_1 \wedge C_2) = CS_s(E, C_1) + CS_s(E, C_2)$

$I_c(E, C_1, C_2)$ implies $CS_c(E, C_1 \wedge C_2) = 1 - (1 - CS_c(E, C_1))\,(1 - CS_c(E, C_2))$

$I_{lr}(E, C_1, C_2)$ implies $CS_{lr}(E, C_1 \wedge C_2) = CS_{lr}(E, C_1)\, CS_{lr}(E, C_2)$

$I_{lr2}(E, C_1, C_2)$ implies $CS_{lr2}(E, C_1 \wedge C_2) = 1 - (1 - CS_{lr2}(E, C_1))\,(1 - CS_{lr2}(E, C_2))$

$I_{ij}(E, C_1, C_2)$ implies $CS_{ij}(E, C_1 \wedge C_2) = 1 - (1 - CS_{ij}(E, C_1))\,(1 - CS_{lr2}(E, C_2))$

It bears remembering, however, that the antecedents are not all mutually satisfiable.[16]

---

[16] For more detailed treatment of the properties of conjunctive causes, see the accompanying notebook at http://fitelson.org/pmcs.nb or http://fitelson.org/pmcs.nb.pdf, pp. 22-30.

References

Cartwright, N. (1979) Causal Laws and Effective Strategies," *Noûs* **13**: 419-437.

Cheng, P. (1997). 'From Covariation to Causation: A Causal Power Theory', *Psychological Review* 104: 367 – 405.

-----, and L. Novick (1990). "A Probabilistic Contrast Model of Causal Induction," *Journal of Personality and Social Psychology* 58: 545 – 567.

Crupi, V., *et al*. (2007). "On Bayesian Measures of Evidential Support: Theoretical and Empirical Issues," *Philosophy of Science* 74(2): 229–252.

Dupré, J. (1984) "Probabilistic Causality Emancipated," in Peter French, Theodore Uehling, Jr., and Howard Wettstein, eds., (1984) *Midwest Studies in Philosophy IX* (Minneapolis: University of Minnesota Press), pp. 169 - 175.

Eells, E. (1991). *Probabilistic Causality*. Cambridge, U.K.: Cambridge University Press.

Eells, E. and Fitelson, B. (2002). "Symmetries and Assymmetries in Evidential Support," *Philosophical Studies* 107: 129–142 <http://fitelson.org/symmetry.pdf>.

Fitelson, B. (1999). "On the Plurality of Bayesian Measures of Confirmation and the Problem of Measure Sensitivity," *Philosophy of Science* 66 (supplement): S362 – S378 <http://fitelson.org/psa.pdf>.

-----. (2001). *Studies in Bayesian Confirmation Theory*. PhD. Dissertation, University of Wisconsin–Madison <http://fitelson.org/thesis.pdf>.

-----. (2008). "A Decision Procedure for Probability Calculus with Applications," *Review of Symbolic Logic* 1: 111–125 <http://fitelson.org/pm.pdf>.

Giere, R. (1979). *Understanding Scientific Reasoning*. New York: Holt, Rinehart and Winston.

Glymour, C. (1998). "Learning Causes: Psychological Explanations of Causal Explanation 1," *Minds and Machines* 8(1): 39-60.

Good, I. J. (1961) "A Causal Calculus I," *British Journal for the Philosophy of Science* **11**: 305-18.

-----. (1962) "A Causal Calculus II," *British Journal for the Philosophy of Science* **12**: 43-51.

Hitchcock, C. (2003) "Causal Generalizations and Good Advice," in H. Kyburg and M. Thalos (eds.)*Probability is the Very Guide of Life* (Chicago: Open Court), pp. 205 - 232.

-----. (2004) "Do All and Only Causes Raise the Probabilities of Effects?" in Collins, John, Ned Hall, and L.A. Paul, eds. *Causation and Counterfactuals* (Cambridge MA: MIT Press), pp. 403 – 418.

Humphreys, P. (1989) *The Chances of Explanation: Causal Explanations in the Social, Medical, and Physical Sciences*. Princeton: Princeton University Press.

Joyce, J. MS. "On the Plurality of Probabilistic Measures of Evidential Relevance," unpublished manuscript.

Lewis, D. (1973). *Counterfactuals*. Cambridge MA: Harvard University Press.

-----. (1976). "Probabilities of Conditionals and Conditional Probabilities," *Philosophical Review* 85: 297 – 315.

-----. (1979). "Counterfactual Dependence and Time's Arrow," *Noûs* 13: 455 – 76.

-----. (1986). "Postscripts to 'Causation'," in *Philosophical Papers, Volume II* (Oxford: Oxford University Press), pp. 173-213.

Lewis, D. (2000). "Causation as Influence," *Journal of Philosophy* 97: 182 - 197.

Mackie. J. (1974). *The Cement of the Universe*. Oxford: Oxford University Press.

Parascondola, M. (1996). "Evidence and Association: Epistemic Confusion in Toxic Tort Law." *Philosophy of Science* 63 (supplement): S168 - S176.

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

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

Skyrms, B. (1980) *Causal Necessity*. New H

aven and London: Yale University Press.

Stalnaker, R. (1968). "A Theory of Conditionals," in N. Rescher, ed., *Studies in Logical Theory*. Oxford: Blackwell.

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

Woodward, J. (1990) "Supervenience and Singular Causal Claims." In D. Knowles, ed. *Explanation and its Limits.* Cambridge, U.K: Cambridge University Press, pp. 211 - 246.

-----. (2003) *Making Things Happen: A Theory of Causal Explanation*. Oxford: Oxford University Press.

Eells:	$CS_e(E, C) = P(E\|C) - P(E\|\sim C)$
Suppes:	$CS_s(E, C) = P(E\|C) - P(E)$
Galton:	$CS_g(E, C) = 4P(C)P(\sim C)[P(E\|C) - P(E\|\sim C)]$
Cheng:	$CS_c(E, C) = (P(E\|C) - P(E\|\sim C))/P(\sim E\|\sim C)$
Lewis ratio:	$CS_{lr}(E, C) = P(E\|C)/P(E\|\sim C)$
	$CS_{lr1}(E, C) = [P(E\|C) - P(E\|\sim C)]/[P(E\|C) + P(E\|\sim C)]$
	$CS_{lr2}(E, C) = [P(E\|C) - P(E\|\sim C)]/P(E\|C)$
Good:	$CS_{ij}(E, C) = P(\sim E\|\sim C)/P(\sim E\|C)$
	$CS_{ij1}(E, C) = [P(\sim E\|\sim C) - P(\sim E\|C)]/[P(\sim E\|\sim C) + P(\sim E\|C)]$
	$CS_{ij2}(E, C) = [P(\sim E\|\sim C) - P(\sim E\|C)]/P(\sim E\|\sim C)$

Table 1: Measures of causal strength

Eells:	$CS_e(E, C) = c + d$
Suppes:	$CS_s(E, C) = c$
Galton:	$CS_g(E, C) = 4cd$
Cheng:	$CS_c\ (E, C) = d/(b + d)$
Lewis ratio:	$CS_{lr}(E, C) = (d + f)/f$
	$CS_{lr1}(E, C) = d/(d + e + f)$
	$CS_{lr2}(E, C) = d/(d + f)$
Good:	$CS_{ij}\ (E, C) = (b + d)/d$
	$CS_{ij1}\ (E, C) = d/(2b + d)$
	$CS_{ij2}\ (E, C) = d/(b + d)$

Table 2: Pictorial Representations

Suppes: $CS_s(E, C) = P(\sim C)CS_e(E, C)$

Galton: $CS_g(E, C) = 4P(C)P(\sim C)CS_e(E, C)$

$= 4P(C)CS_s(E, C)$

Cheng: $CS_c(E, C) = CS_e(E, C)/P(\sim E|\sim C)$

$= CS_s(E, C)/P(\sim E \wedge \sim C)$

$= CS_g(E, C)/4P(C)P(\sim E \wedge \sim C)$

Lewis ratio: $CS_{lr1}(E, C) = [CS_{lr}(E, C) - 1]/[CS_{lr}(E, C) + 1]$

$=CS_e(E, C)/[P(E|C) + P(E|\sim C)]$

$CS_{lr2}(E, C) =1 - 1/CS_{lr}(E, C)$

$= CS_e(E, C)/P(E|C)$

$= CS_s(E, C)/P(E|C)P(\sim C)$

$= CS_g(E, C)/4P(E \wedge C)P(\sim C)$

$= CS_c(E, C)[P(\sim E|\sim C)/P(E|C)]$

Good: $CS_{ij}(E, C) = CS_{lr}(\sim E, \sim C)$

$CS_{ij1}(E, C) = [CS_{ij}(E, C) - 1]/[CS_{ij}(E, C) + 1]$

$=CS_{lr1}(\sim E, \sim C)$

$= CS_e(E, C)/[P(\sim E|C) + P(\sim E|\sim C)]$

$CS_{ij2}(E, C) = 1 - 1/CS_{ij}(E, C)$

$= CS_c(E, C)$

$= CS_e(E, C)/P(\sim E|\sim C)$

$= CS_s(E, C)/P(\sim E \wedge \sim C)$

$= CS_g(E, C)/4P(C)P(\sim E \wedge \sim C)$

$= CS_{lr2}(\sim E, \sim C)$

Table 3: Inter-definability of the Measures

	Eells	Suppes	Galton	Cheng	Lewis Ratio	Good
Eells	G-E	II-E	II-E	None	None	None
Suppes	II-E	G-E	II-E	None	None	None
Galton	II-E	II-E	G-E	None	None	None
Cheng	None	None	None	G-E	None	G-E
Lewis Ratio	None	None	None	None	G-E	None
Good	None	None	None	G-E	None	G-E

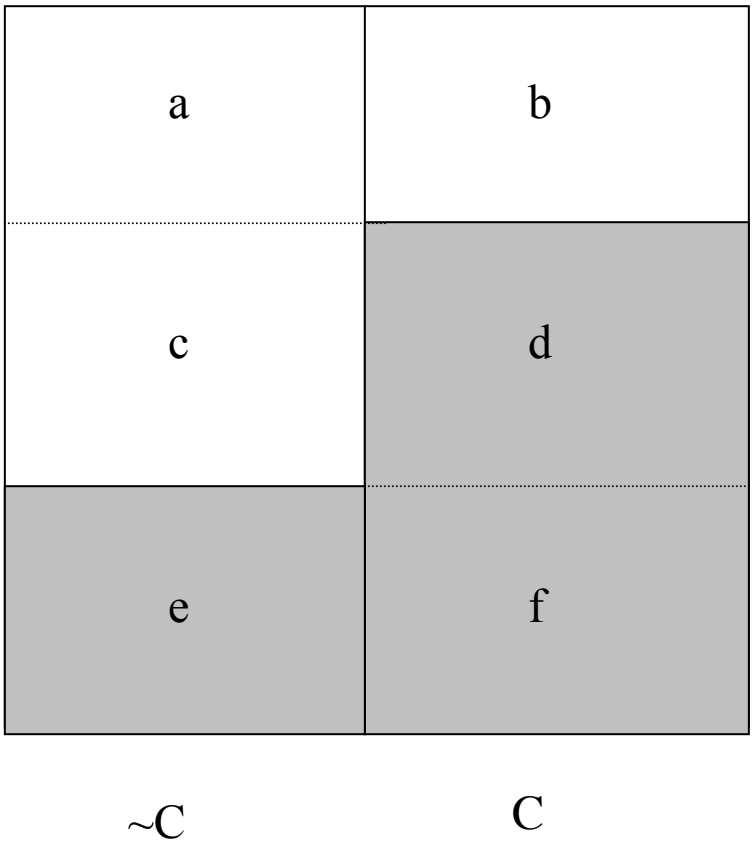
Table 4: Ordinal Equivalences Between Measures

	(CPC)	(COC)	(CPO)
Eells	Yes	Yes	Yes
Suppes	Yes	No	No
Galton	Yes	Yes	Yes
Cheng	No	No	No
Lewis Ratio rescaling #1 CS_{lr1}	No	Yes	No
Lewis Ratio rescaling #2 CS_{lr2}	No	No	No
Good rescaling #1 CS_{ij1}	No	Yes	No
Good rescaling #2 CS_{ij2}	No	No	No

Table 5: Continuity Properties of Measures

	Eells	Suppes	Galton	Cheng	Lewis Ratio	Good
Eells	All	All	All	None	None	None
Suppes	All	All	All	None	None	None
Galton	All	All	All	None	None	None
Cheng	None	None	None	All	None	All
Lewis Ratio	None	None	None	None	All	None
Good	None	None	None	All	None	All

Table 6: Do measures C_1 and C_2 agree on All, Some, or None of their I_{CS}-judgments?



E =

Figure 1

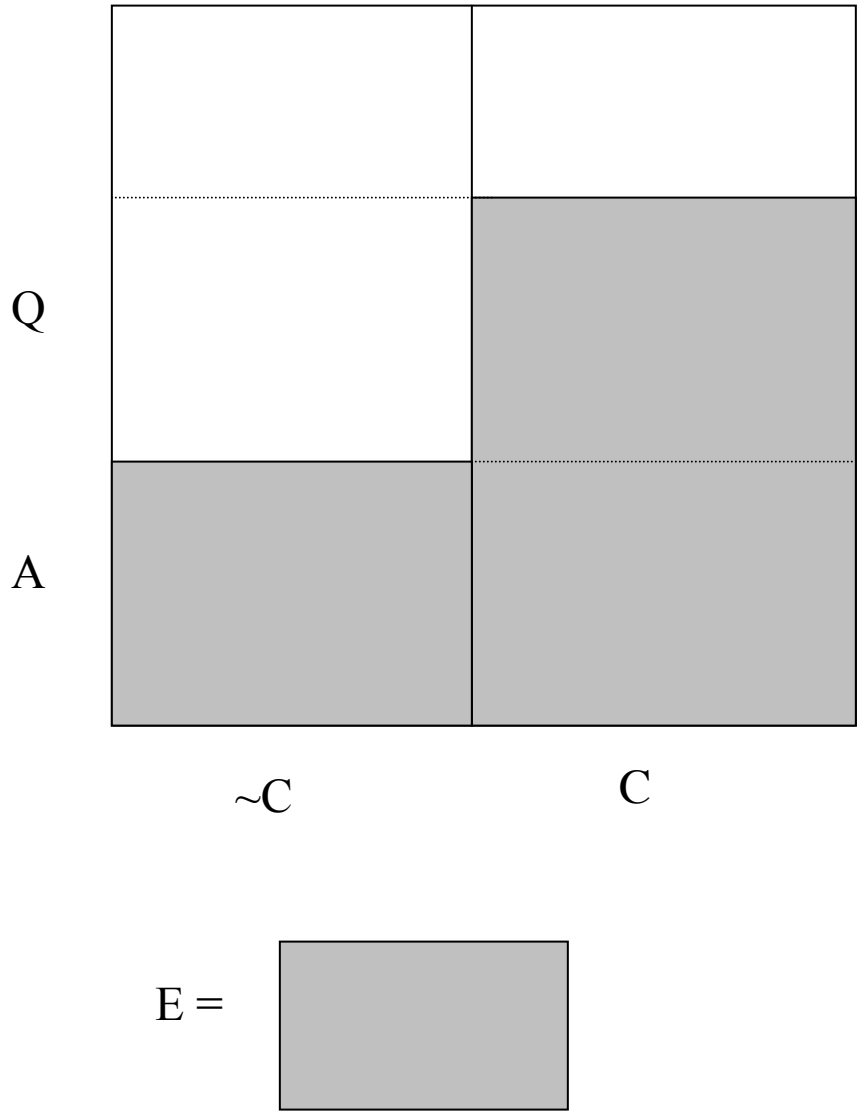


Figure 2

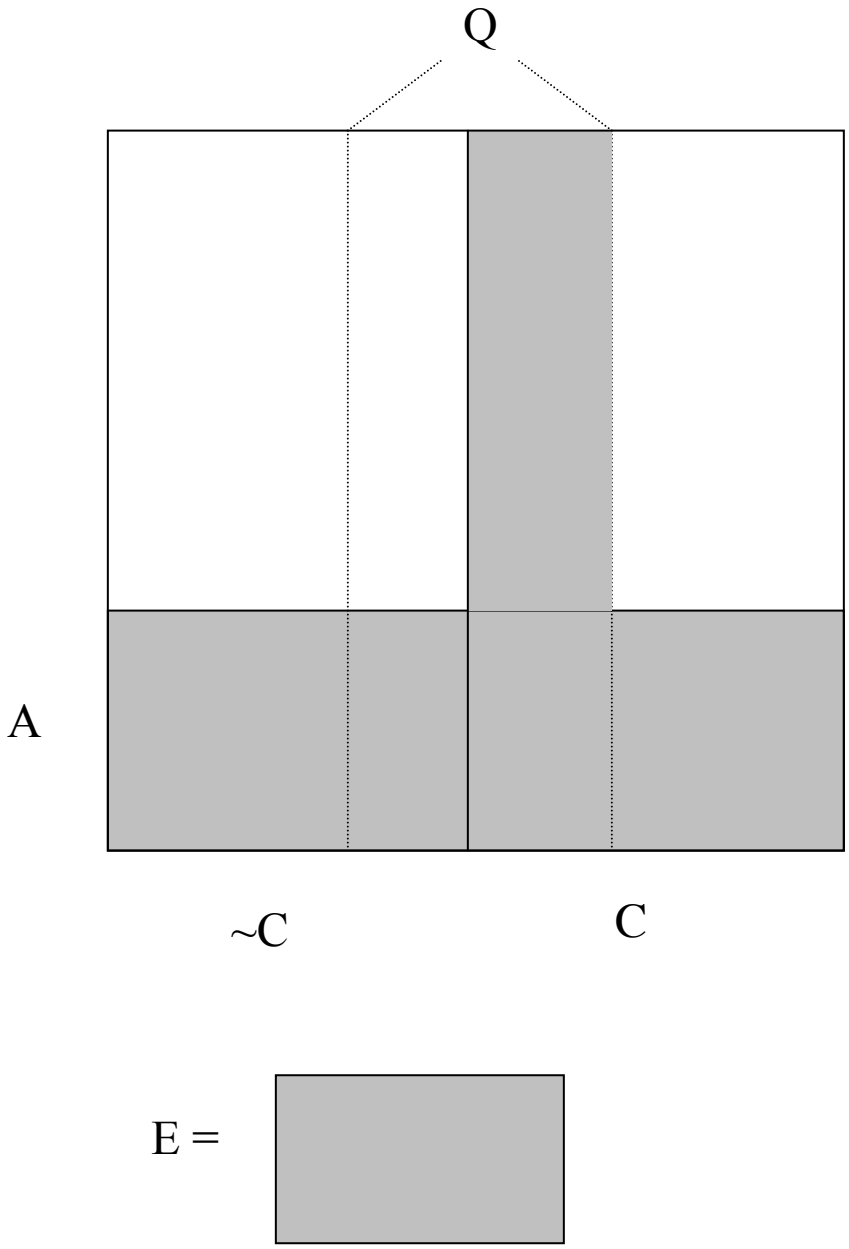


Figure 3