The pauper's problem:

Chance, foreknowledge and causal decision theory*

Adam Bales

September 14, 2015

Abstract

In a letter to Wlodek Rabinowicz (Lewis, 1982), David Lewis introduced a decision scenario that he described as "much more problematic for decision theory than the Newcomb Problems". This scenario, which involves an agent with foreknowledge of the outcome of some chance process, has received little subsequent attention. However, in one of the small number of discussions of such cases, Price (2012) has argued that cases of this sort pose serious problems for causal decision theory (the version of decision theory championed by Lewis and many others). In this paper, I will argue that these problems can be overcome: scenarios of this sort do not pose fatal problems for this theory as there are versions of CDT that reason appropriately in these cases. However, I will also argue that such cases push us toward a particular version of CDT developed in Rabinowicz, 2009.

*This is a penultimate version of a published paper. The final publication is available at Springer via http://dx.doi.org/10.1007/s11098-015-0560-8. For assistance with writing this paper, thanks to Arif Ahmed, Bill D'Alessandro, Richard Holton, Shay Logan, Huw Price, Wlodek Rabinowicz, Shyane Siriwardena, Georgie Statham, Brian Weatherson and an anonymous reviewer for Philosophical Studies. For related discussion, thanks to Andy Egan and Chris Hitchcock.

Consider a puzzle that we could call the pauper's problem (quoted from Lewis, 1982):[1]

> You know that spending all you have on armour would greatly increase your chances of surviving the coming battle, but leave you a pauper if you do survive; but you also know, by news from the future that you have excellent reason to trust, that you will survive. (The news doesn't say whether you will have bought the armour.) Now: is it rational to buy the armour?

In an earlier paper (his 1981, p. 18), Lewis described cases of this sort (by which he means cases where an agent has foreknowledge of the output of a chance process), as "much more problematic for decision theory than the Newcomb Problems".[2] Yet while Newcomb's problem has inspired a vast literature, the pauper's problem and related cases have been followed up in just three papers (Rabinowicz, 2009, Price, 2012 and Hitchcock, 2015). In this paper, I aim to partially undo this neglect by considering whether these cases provide counterexamples to causal decision theory (CDT), a prominent theory of choice. This theory counts Lewis among its proponents and when he described the pauper's problem as problematic for decision theory, it was surely CDT that he had in mind. There are grounds, then, for concern on behalf of the proponent of CDT. However, I will show that on further reflection these grounds disappear. In particular, I will show that: (a) CDT survives the challenge raised by the pauper's problem and related cases (including those in Price, 2012); and (2) these cases nevertheless push us toward a particular version of CDT (which I will call Rabinowiczian CDT).

[1] For both Lewis, 1982 and Rabinowicz, 1982a I will cite the original letters, which are reproduced in a draft of Price, 2012 (available at http://philsci-archive.pitt.edu/4894/1/ChewcombWithAppendix.pdf).

[2] The foreknowledge cases originated in Sobel, 1980. Newcomb's problem is a well-known decision scenario to be discussed below.

1 Clarifying the problem

As a first step to evaluating the impact of the pauper's problem, it will help to fill in the details of this scenario. First, note that what it is rational for the pauper to do depends, in part, on how undesirable she takes death and poverty to be (for example, the case is dramatically different if she desires, rather than fears, death). I will presume (as is standard in decision-theoretic discussions) that these desires can be represented by numerical utilities, where a higher number represents a more desirable outcome. So, in the pauper's problem, we can (following Rabinowicz, 1982a) value survival at 100 and impoverishment at -10 and sum these numbers to determine each possible outcome's utility (for example, the utility of an outcome where the pauper survives but is impoverished is 90). These utilities are outlined in Table 1, which takes the world to occupy one of three states: (1) in the full-blow-state, the pauper suffers an attack that is lethal regardless of whether armour is worn; (2) in the glancing-blow-state, she suffers an attack that is lethal only without armour; and (3) in the dodge-state she dodges all attacks.[3]

	Full Blow	**Glancing Blow**	**Dodge**
Buy Armour	-10	90	90
Refrain	0	0	100

Table 1: Utilities in the pauper's problem.

A further point: as Lewis intended the case, the pauper's survival (or

[3]Why not just a survival and a death state? Because the pauper's decision causally influences her survival and this would complicate discussion. Note an assumption: wearing armour won't change the pauper's chance of being hit (by, for example, making her reckless). Another assumption: the agent prefers not to be impoverished even in death (as Rabinowicz notes in his letter to Lewis, perhaps she has family that she desires to leave money to if she dies).

death) is determined by a non-trivial chance process (Lewis was concerned with what he called *abnormal* cases: cases where an agent has foreknowledge of the outcome of some *chance* process).[4] We can stipulate that this chance process outputs full blow with probability 0.3, glancing blow with probability 0.3 and dodge with probability 0.4. This is the pauper's problem in the form that I will be discussing.

2 A counterexample?

The most obvious way that the pauper's problem could be problematic for CDT would be if it were a counterexample to this theory (that is, if CDT gave the wrong guidance in this case). However, a difficulty arises with establishing this: Lewis, 1982 claims that there are compelling arguments for various distinct positions in this case and if Lewis is right about this it will be hard to determine what guidance is appropriate in this case and hence hard to determine whether the case is a counterexample. Let's consider what these arguments might be.

2.1 Argument 1. Why refrain?

A simple argument for refraining notes that the pauper knows that: (1) if she buys armour, she will survive and be impoverished (and will have caused her impoverishment); and (2) that if she doesn't buy armour, she will survive and not be impoverished (and will have caused herself to not be impoverished). As the second of these outcomes is better than the first, it might seem irrational to buy armour.

[4]The chanciness makes the pauper's problem interestingly distinct from the cases in Egan, 2007 (as CDT treats chancy cases interestingly differently to other cases).

2.2 Argument 2. Why buy?

A different argument supports buying armour: (1) rational choice (it might be argued) is not about providing ourselves with good news but about having the best causal effect (for example, taking an offered $1000 is rational not because doing so provides us with the good news that we have $1000 but because it *causes* us to gain $1000); (2) while buying armour doesn't provide good news (because the pauper knows she will survive), it does have a good causal effect (it increases her objective chance of survival from 40% to 70%); and therefore (3) it is rational to buy armour.

2.3 Argument 3. Why vary your behaviour?

Finally, according to a third argument (drawn from Rabinowicz, 1982a), what the pauper ought to do depends on what she believes she will do.[5] Consider the extreme cases. First, if the pauper is certain that she won't buy armour then her knowledge that she will survive entails that the dodge state holds (as otherwise she would die without armour). If so then the pauper will survive regardless of whether she wears armour and so would be foolish to spend money on armour. So if the pauper is sure that she won't buy armour, she is rational to not do so. On the other hand, if she is certain that she will buy armour then her certainty that she will survive entails that either the dodge state or the glancing-blow state holds (as she would die if the full-blow state held). Presuming her credence (that is, subjective probabilities) in these states is proportional to their chances, she will have a credence of $\frac{3}{7}$ in the glancing-blow state (and hence this same credence that armour would save her life). Consequently, she ought to buy armour (given

[5]This argument presumes, contra for example Price, 2012, pp. 528–531, that a deliberating agent is able to have credences regarding her own actions.

how much she values survival and how little she cares about impoverishment, armour is worth the investment). So if the pauper is certain that she will buy armour then she ought to do so. Given this, what the pauper ought to do depends on her beliefs about what she will do. Further, this same reasoning can be extended to cases where the agent is uncertain about her future behaviour. Once the calculations are carried out, it turns out that this reasoning entails that if the pauper has a credence of about 0.233 that she will buy then buying and refraining are both rationally permissible but that if this credence is lower then refraining will be rationally required (and if it is higher, then buying will be). In the non-extreme cases too, what the pauper ought to do depends on her beliefs about what she will do.

There are arguments, then, for three different positions in the pauper's problem. Presuming we take each of these to have some force, Lewis is right that there are difficulties with determining what decision is rational in this case and hence difficulties with determining whether it is a counterexample to CDT.

3 Causal decision theory

At this stage, some will be tempted to commit to one of the above positions, perhaps on the basis of intuition, so as to determine whether the scenario is a counterexample to CDT. However, relying on controversial intuitions in strange cases strikes me as a dangerously overused move so I will instead champion a different approach.

First, let's clarify the target: speaking loosely, according to CDT the rational decision is that with the best expected causal effects. This theory is then formalised in various ways, such that these different formalisations do not always provide the same guidance. Given this, a solution suggests itself:

if it could be shown that, for each plausible take on the pauper's problem, some version of CDT gives the desired guidance, then it follows that the pauper's problem is not a general counterexample to CDT (as regardless of which position on the case is correct some version of CDT survives). I proceed, then, to establish this claim by discussing three versions of CDT.

3.1 Lewisian CDT

The version of CDT defended in Lewis, 1981 endorses the rational requirement of refraining from buying armour.[6] In informal terms, this can be established by noting that Lewisian CDT reduces to evidential decision theory (EDT), an alternative theory of choice, in cases like the pauper's problem (where the agent is certain about the causal influence of her decisions on the chances: see Lewis, 1981, p. 11). Now according to EDT, the rational decision is that which provides the best evidence and in the pauper's problem this will be the decision to refrain. After all, the agent knows that she will survive so the only evidence the decisions provide her is evidence about whether she will be impoverished or not (and refraining provides the more desirable evidence that she will not be). So EDT, and therefore Lewisian CDT, will endorse refraining in the pauper's problem.

Giving this same argument in formal terms requires some background. First, let a dependency hypothesis be a, "maximally-specific proposition about how the things [the agent] cares about do and do not depend causally on his present actions" (Lewis, 1981, p. 11). In chancy cases, these can be thought of as having the form: "if I were to make this decision, the

[6]Lewis himself set abnormal cases aside when developing CDT (Lewis, 1981, p. 18). Perhaps this was merely a simplifying assumption or perhaps decision theory should not apply in these cases (see Lewis, 1982). However, I set aside this second option (which would push us toward a restricted CDT that didn't apply to abnormal cases) as it seems that guidance should be provided in some abnormal cases (see §§4–6).

chances on which the things I care about depend would be thus-and-so, whereas if I were to make this alternative decision, they would be so-and-thus, whereas...". In other words, dependency hypotheses can be thought of as conjunctions of counterfactuals with consequents stated in terms of chances, with one counterfactual for each possible decision.[7] In symbols, then, the only dependency hypothesis in the pauper's problem is the following (where $A \boxdot\!\!\rightarrow B = p, C = q$ is the counterfactual, "if A were the case, the chance of B would be p and the chance of C would be q):[8]

K_1: $(B \boxdot\!\!\rightarrow G = F = 0.3, D = 0.4) \wedge (R \boxdot\!\!\rightarrow G = F = 0.3, D = 0.4)$

In other words, according to K_1 the chance of a glancing-blow (G) and a full-blow (F) is 0.3 and the chance of dodging (D) is 0.4, regardless of whether the agent decides to buy (B) or refrain (R).

With this in mind, and letting K range over the dependency hypotheses, Lewisian CDT works by calculating, for each decision A, an expected utility (EU) as follows:

$$EU(A) = \sum_K Cr(K)V(KA)$$

Here, $Cr(K)$ is the agent's credence that the dependency hypothesis is true and $V(KA)$ represents the desirability of this hypothesis being true if A is decided upon. So a decision's EU is a sum of the values of the outcomes

[7] These must be non-backtracking counterfactuals, which we can think of as meaning roughly that they hold fixed the past until shortly before the decision (see Lewis, 1979).

[8] It is crucial that K_1 be the sole dependency hypothesis here (rather than, for example, there being three hypotheses corresponding to the possible chance-process outputs). Firstly, this is key to adequately representing Lewis's views (other Lewisian dependency hypotheses will misdescribe the chances in the pauper's problem). Second, getting this right is key to determining Lewisian CDT's guidance (briefly: this is crucial because Lewisian CDT treats the evidence provided by decisions differently depending on whether it shifts the agent's credences between worlds within a hypothesis or between worlds in different hypotheses and so the boundary between these categories is key).

that might result from the decision, weighted by the agent's credence that the world is such that the decision will lead to those outcomes. A decision is then rational if it maximises EU (that is, if it has an EU at least as high as any alternative).

Given one more detail, it will be possible to see why this theory reduces to EDT in the pauper's problem. In particular, given that value is assigned in the first instance to possible worlds (for example, a world in which the pauper survives and buys armour is assigned a value of 90), it remains to be said how we should calculate $V(KA)$, given that KA is a proposition (set of worlds) rather than a world itself. According to Lewisian CDT, this value is calculated as follows (where W ranges across the worlds):

$$V(KA) = \sum_W Cr(W \mid KA)V(W)$$

So the value of a KA is a weighted sum of the values of the worlds, where the weighting is the agent's credence that that world is the actual world after conditioning on the assumption that KA holds. With this in hand, we can note what happens when the agent concentrates all her credence in one dependency hypothesis (and hence assigns this hypothesis a credence of 1). Letting K_n be this dependency hypothesis, the EU of a decision, A, can be calculated as follows:[9]

[9](2) comes from (1) as the agent has credence 1 in K_n and credence 0 in all other dependency hypotheses; (3) comes from (2) by substituting in the full formula for $V(KA)$; and (4) comes from (3) because, given that the agent already has credence 1 in K_n worlds, conditionalising on AK_n is equivalent to conditionalising on A.

$$EU(A) = \sum_K Cr(K)V(KA) \tag{1}$$

$$= 1 * V(K_nA) \tag{2}$$

$$= \sum_W Cr(W \mid K_nA)V(W) \tag{3}$$

$$= \sum_W Cr(W \mid A)V(W) \tag{4}$$

Equation (4) is the equation used in EDT and so on the assumption that the agent concentrates all of her credence in one dependency hypothesis, Lewisian CDT reduces to EDT.

Now return to the pauper's problem. Here, the agent has credence in just one dependency hypothesis (K_1) and so Lewisian CDT calculates EUs according to equation (4) (that is, according to EDT). With this in mind, consider first the EU of buying. The only time the agent will have positive credence in a world after conditioning on buying is when that world is: (a) one in which she survives (because she is certain she will survive, all of the pauper's initial credence is in survival worlds and conditionalising simply rearranges her credences within this set of worlds and so cannot lead her to have positive credence in non-survival worlds); and (b) one in which she is impoverished (because in all the worlds where she buys armour she is impoverished). All such worlds will have a value of 90 (100 for survival, -10 for impoverishment). As such, a credence-weighted sum of these values must also be 90 (given that the credences across all worlds must sum to 1) and so the EU of buying will be 90. Now consider refraining. After conditioning on refraining, the agent will have positive credence in worlds in which she: (a) survives; and (b) is *not* impoverished (as she is never impoverished when

she refuses armour). All such worlds have a value of 100 and so refraining has an EU of 100. Consequently Lewisian CDT will endorse the rational requirement of refraining.

3.2 Sobelian CDT

Contra Lewisian CDT, the version of CDT in Sobel, 1980 will endorse the rational requirement of buying armour. Informally, it's easy to see why: Sobelian CDT focuses on what the chances would be if each decision were made and the pauper's knowledge about her survival is irrelevant to these chances (and so is ignored by Sobelian CDT).[10] As such, this theory endorses buying armour as it has a 30% chance of saving the pauper (which is worth $0.3 * 100 = 30$) and this outweighs the cost of certain impoverishment (which is worth -10).

Formally, Sobelian CDT appeals to a tendency function that, in the first instance, is applied to worlds, such that $W_A(W')$ specifies the chance, at W, of W' being the case if A were the case (I will describe the application of this function as *imaging on* A, as the tendency function is an imaging function).[11] For example, if A were the proposition that some fair coin was tossed (such that this coin *was* tossed at W) and W' is the world that would eventuate if this coin came up heads, then $W_A(W') = 0.5$. In Sobelian CDT this tendency function on worlds is taken as primitive and is then used to define a tendency function on credence functions, such that $Cr_A(W') = \sum_W Cr(W)W_A(W')$. Less formally: this tendency function on credence functions starts with the

[10] In contrast, in appealing to $Cr(W \mid KA)$—in calculating $V(KA)$—Lewisian CDT becomes responsive to inadmissible evidence (in the sense of Lewis, 1980) and so need not follow the chances in cases involving such evidence. In relation to the armour case, and other abnormal cases, this is the key technical difference between these two versions of CDT.

[11] An imaging function is a function from a world and a proposition to a probability distribution over worlds such that only worlds in which the proposition holds receive positive probability.

worlds in which the agent has positive credence and then redistributes her credences in each world in accordance with the chances at that world (or, more accurately, with what the chances would be if the proposition being imaged on were true). With this notation in hand, Sobelian CDT then defines the EU of a decision as follows:[12]

$$EU(A) = \sum_W Cr_A(W)V(W)$$

We can now apply this theory to the pauper's problem. In this case, the pauper is certain about the chances (that is, certain that the glancing-blow and full-blow states each have a chance of 0.3 and that the dodge state has a chance of 0.4). As such, these will be the chances at all worlds in which the pauper has positive credence and so (given that these chances are independent of the pauper's decision) once the tendency function is applied the agent's credence will divide between the glancing-blow, full-blow and dodging worlds in proportion to these chances, regardless of which decision is imaged on (because the tendency function redistributes the agent's credences in line with what she believes the chances to be). Further, as imaging on A picks out only A worlds, $Cr_B(W)$ (that is, credence in W after imaging on buying armour) will be positive only if W is an armour-buying world. This means that the values of the worlds relevant to assessing the EU of buying can be read off the buying row of Table 1 (for example, the value of the dodge worlds is found in the rightmost cell of this row). It follows then, that $EU(B) = (0.3 * -10) + (0.3 * 90) + (0.4 * 90) = 60$. Similarly, $Cr_R(W)$ (that is, credence in W after imaging on refraining from buying armour) will be positive only if W is a refraining world and so the values of the worlds relevant to assessing the EU of refraining can be read off the refraining row

[12]This formula requires the assumption that $W_A(W')$ is defined for all worlds, W and W', and all actions, A (see Rabinowicz, 1982b, pp. 308–309 for a discussion of how to weaken this assumption).

of the table. It follows then that $EU(R) = (0.3 * 0) + (0.3 * 0) + (0.4 * 100) = 40$. As 60 is higher than 40, Sobelian CDT labels buying armour as rationally required. We now have two versions of CDT that deliver two different solutions in the pauper's problem.

3.3 Rabinowiczian CDT

Consider a fair (chancy) coin that will shortly be tossed. According to our ordinary notion of chance, the chance of this coin coming up heads is 0.5. These ordinary chances underpin the tendency function of Sobelian CDT. In contrast to this, Rabinowicz (1982a; 2009) considers a modified version of CDT where the tendency function is underpinned by a distinct notion of chance (call this the outcome-responsive chance of heads).[13] According to the outcome-responsive chances, the chance, before the toss, of the coin coming up heads is 1 if the coin will actually end up coming up heads and is 0 otherwise.[14] In more detail: on this view the tendencies will be trivial (0 or 1) when considering: (1) actual circumstances (that is, if you will actually bet heads then the tendency of heads after imaging on betting heads will be 1, if the coin will come up heads, and 0 otherwise); and (2) counterfactual circumstances *if the chance process is causally independent of the factor being imaged on* (that is, if you will actually bet heads then the tendency of heads should still be 1 or 0, as appropriate based on how the coin will land, even after imaging on betting *tails*, as long as the outcome of the coin toss is independent of the bet made).[15] Call a version of CDT that

[13] Rabinowicz talks instead of *centered* chances.

[14] Are "outcome-responsive chances" chances in anything but name? They are *probabilities* that can take non-trivial values (see Rabinowicz, 2009) and which play something close to chance's epistemic role as per Lewis, 1980. Regardless, little rests on this issue (the key question is not whether Rabinowiczian CDT relies on chances but whether it is a desirable theory of choice).

[15] More formally: (1) says that $W_A(W) = 1$ if W is an A-world; and (2) says that, if W is an A-world then $W_{\neg A}(W') > 0$ only if W' is a world in which the chance process has

relies on a tendency function of this sort *Rabinowiczian*. In relation to the pauper's problem, Rabinowiczian CDT entails that what the pauper ought to do depends on what she believes she will do.[16]

Take two sample cases. First, if the agent is certain that she will refrain from buying armour then she knows that the chance process will terminate in the dodge-state (as she knows she will survive and will do so only if she dodges, given that she refrains from buying). So she will have positive credence only in worlds where the chance process outputs dodging. With this in mind, we can evaluate the decision to refrain from buying armour. The agent is certain that she will make this decision and so images on a proposition (that she will refrain) that she is certain is actual. As such, (1) entails that the Rabinowiczian tendency function will assign dodging a chance of 1 (as the agent is also certain that this will be the actual output of the chance process). Given that dodging worlds are valued at 100 if the agent refrains from buying armour, this means that Rabinowiczian CDT will calculate the EU of refraining as $1 * 100 = 100$. On the other hand, consider the decision to buy armour (still from the perspective of the agent who is certain she will refrain from doing so). The agent believes that this is a counterfactual circumstance but (we may stipulate) also believes that the chance process is independent of her decision and so, by (2), after imaging the tendency of the chance process outputting dodging is 1. Given that dodging worlds are valued at 90 if the agent buys armour, this means that Rabinowiczian CDT will calculate the EU of buying as $1 * 90 = 90$. Given

the same output as it does in W (in cases where the chance process is independent of A). This specifies a *family* of functions because it leaves open what the tendencies should be in counterfactual circumstances when the chance process *is* dependent on the proposition imaged on (as, for example, when a coin will only be tossed if the proposition holds).

[16]My calculations differ from Rabinowicz's (in his 1982a; 2009) because: (a) we use different chances in outlining the pauper's problem; and (b) Rabinowicz doesn't assume, as I do, that the chance process's output is independent of the decision (see fn 3). My assumption here is unrealistic but simplifies things without undermining my results.

the EU of each decision, if the agent is certain that she will refrain then Rabinowiczian CDT will label refraining as rationally required.

On the other hand, if the pauper is certain that she will buy then she is certain that the chance process terminates in either the glancing-blow or the dodge state (as she knows she will survive and knows she would die if the process terminated in a full blow). Presuming she divides her credence between these two possibilities in proportion to their original chances (as she will if her prior credences follow the chances and she conditionalises on survival) the pauper will have a credence of $\frac{3}{7}$ in glancing-blow worlds (that is, worlds in which the chance process outputs a glancing blow) and a credence of $\frac{4}{7}$ in dodge worlds. Now consider the decision to buy armour. Because the agent is certain that she will buy, she will have positive credence only in worlds in which she does so. As such, by (1), imaging on buying armour will move all her credence in glancing-blow worlds to glancing-blow worlds and all of her credence in dodging worlds to dodging worlds. It follows that she will retain her credence in each of these possibilities and so that the EU of buying will be $(\frac{3}{7} * 90) + (\frac{4}{7} * 90) = 90$. On the other hand, consider the decision to refrain from buying armour. By (2), imaging on this choice will also leave the credences as they were before imaging and so the EU of refraining would be $(\frac{3}{7} * 0) + (\frac{4}{7} * 100) = 57$. Given these credences, the theory will label buying as rationally required (as $90 > 57$). Consequently, Rabinowiczian CDT entails that the agent who is sure that she will refrain ought to refrain and the agent who is sure that she will buy ought to buy. As such, according to this theory what the pauper ought to do depends on her credences regarding what she will do.[17]

We now have versions of CDT that correspond to each of the three

[17]This point also extends to cases where the agent is uncertain how she will decide.

positions that I previously argued one could take on the pauper's problem. Even without determining which of these positions is correct, then, we can conclude that this case fails to provide a general counterexample to CDT (as it fails to show that we must abandon CDT in all of its forms).[18] Given my unwillingness to commit to a view regarding what is rational in the pauper's problem, this is as far as I will get in discussing this scenario.

4 The Chewcomb problems

CDT survives the pauper's problem. Does it survive other abnormal cases? In the remainder of this paper, I will consider this questions by reflecting on a series of cases from Price, 2012, which may seem like promising candidates to be general counterexamples to CDT.[19]

As a first step to outlining these cases, consider two principles. First:

> EVIDENTIAL IRRELEVANCE: The evidence that would be provided by an agent's decisions is not directly relevant to the rationality of those decisions (at least insofar as this evidence does not result from the decision's causal influence).

The proponent of CDT appears likely to accept this principle (at least if it's interpreted carefully).[20] Lewis himself, for example, states explicitly

[18]Though it remains possible that the pauper's problem is part of a set of cases, such that no theory of CDT reasons correctly in all scenarios in the set. This possibility is worth considering but I set it aside here.

[19]Price's paper contains many threads, most of which I set aside: my interest is solely in whether Price's cases are novel counterexamples to CDT.

[20]The word "directly" is key. CDT's proponents might think the evidence provided by decisions relevant to: (1) the values in the EU formula (Lewisian CDT); or (2) updating the agent's unconditional credence prior to decision (as in Rabinowiczian CDT which uses such evidence plus the agent's beliefs about how she will decide to update her credences). What this proponent denies is that this evidence has further relevance. Further, even with this in mind, it might be possible to deny that CDT is committed to EVIDENTIAL IRRELEVANCE (this will depend on how exactly one defines CDT and how one spells out

that such evidence needs to be excluded from consideration when discussing rational choice, speaking of, "that news-bearing aspect of the options that we meant [in developing CDT] to suppress" (where this "news-bearing aspect" is precisely the non-causal evidence that would be provided by the agent's decisions). More generally, it seems likely that proponents of all standard versions of CDT will have to accept this principle. After all, consider the well-known Newcomb's problem (Nozick, 1969), which can be outlined as follows:

> A perfect predictor of your behaviour presents you with two boxes, one opaque and one transparent, and offers you the choice of either taking both boxes (called two-boxing) or of taking just the opaque box (one-boxing). The transparent box contains $1000 and, you are told, the opaque box was filled based on a prediction of your behaviour: it contains nothing if you were predicted to two-box and $1,000,000 if you were predicted to one-box.

Standard versions of CDT are designed to entail the rational requirement of two-boxing in Newcomb's problem and they do so precisely because they ignore the evidence that the agent's decision provides. That is, they ignore the fact that one-boxing provides evidence that the opaque box contains $1,000,000 and the fact that two-boxing provides evidence that it is empty. Given that these theories ignore this evidence in Newcomb's problem, it's natural to think that they must ignore it generally (or otherwise the proponent of the theory may struggle to justify why such evidence can be ignored

EVIDENTIAL IRRELEVANCE). However, rather than engaging further with these issues, I will simply assume that CDT is committed to EVIDENTIAL IRRELEVANCE for now (while outlining Price's challenge) and will then return to some of these complexities later.

in some cases but not others). As such, it's natural to think that proponents of all standard versions of CDT must accept EVIDENTIAL IRRELEVANCE.

Now turn to a second principle. To do so, consider the connection between rational credences about the outputs of chance processes and the chances themselves. According to Lewis (1980), these credences should generally follow the chances (or, more accurately, the agent's beliefs about the chances). So if a coin has a chance of 0.5 of coming up heads (and the agent knows this) then it is typically rational to have a credence of 0.5 that the coin will come up heads. However, now imagine cases where the agent somehow gains advanced knowledge of the output of the chance process (perhaps via a crystal ball). In such cases, Lewis says that rational credences should reflect this evidence (which he calls inadmissible evidence but which I will call *prophetic evidence*) rather than continuing to track the chances.[21] The proponent of CDT tends to agree and so accepts:

> PROPHETIC RELEVANCE: An agent's credences should account for prophetic evidence.

The basic insight behind Price's cases is that while the proponent of CDT seems to be committed to both PROPHETIC RELEVANCE and EVIDENTIAL IRRELEVANCE, these principles can be brought into tension.[22] To see this, first consider *unconditional Chewcomb*:

> God offers you two bets on the outcome of a fair coin toss (and you must take exactly one of these). A bet on heads pays $100

[21] I depart from Lewis's terminology because *inadmissible* evidence is precisely the evidence that *can* influence credences and so is, in a sense, admissible in the current context.

[22] Some might question this apparent tension: prophetic relevance relates to epistemic rationality and evidential irrelevance to practical rationality (and it is hardly news that these two might clash). Response: credences play an important role in practical rationality and an agent's credences are plausibly determined by her hypothetical behaviour: these links make room for the tension (though see fn 29).

> if the coin comes up heads and \$0 otherwise (the bet has no cost). A bet on tails pays \$50 if the coin comes up tails and \$0 otherwise. However, while the coin is fair, Satan informs you that, as a matter of actual fact, on occasions where you bet, the coin comes up tails 99% of the time.

Presuming we trust Satan, this information is prophetic evidence (that is, it is evidence about the actual outcome of a chance process). Consequently, by PROPHETIC RELEVANCE, this information should be taken into account and so you should have a credence of 0.99 that the coin will come up tails. With this credence in hand, it is clearly rational to bet on tails (as you are almost certain that the coin will come up tails). So far, so good (no problem for CDT yet arises because EVIDENTIAL IRRELEVANCE doesn't come into play in this case: you know Satan's information in advance of your decision and so this evidence isn't due to the decision).[23]

Now Price introduces a second case, *conditional Chewcomb*, which is identical to unconditional Chewcomb except: (a) you have a third option of refraining from betting; and (b) while Satan provides his information before you make a decision, he tells you only that the coin comes up heads 99% of the time *on those occasions where you make some bet* and says nothing about those occasions where you refrain. In this case, Satan's evidence does result from the agent's decision (as this evidence is only available if you bet) and so EVIDENTIAL IRRELEVANCE entails that this evidence should be ignored.[24] Given this, the agent should retain her chance-following credence of

[23]The scenario's coherence or plausibility might be doubted but I set this possibility aside (and show that CDT survives anyway).

[24]In fact, if the agent has credences regarding her decisions then accounting for Satan's evidence need not violate plausible versions of EVIDENTIAL IRRELEVANCE. Further, in such circumstances a number of the claims made in this paragraph will be false (in particular, the agent's credences won't follow the chances and the agent will gain some evidence from her decision regardless of how she decides). For now, then, I simply note that even with

0.5 in each of heads and tails and, given these credences, it is rational to bet heads (as a winning heads bet has a better payout than a winning tails bet). Again, conditional Chewcomb does not, by itself, cause problems for CDT: here EVIDENTIAL IRRELEVANCE comes into play but PROPHETIC RELEVANCE does not (at least not in a clear manner): in this case the agent lacks the prophetic evidence until she has decided and so lacks it at when she deliberates.[25] This gives the proponent of CDT room to deny that this evidence should be taken into account at the time of decision. As with unconditional Chewcomb, then, conditional Chewcomb does not by itself undermine CDT.

Still, a trap can now be sprung: clearly, Price says, the same decision must be rational in the two cases (because adding in a clearly flawed option shouldn't change what is rational in a scenario). Yet EVIDENTIAL IRRELEVANCE and PROPHETIC RELEVANCE jointly entail that different decisions are rational in the two cases and so at least one of these principles has to go. Price thinks PROPHETIC RELEVANCE is plausible and so we ought reject EVIDENTIAL IRRELEVANCE and hence reject CDT, which we are assuming is committed to this principle. CDT, then, faces a serious challenge.

5 Lewis and Chewcomb

As a first step to reflecting on this challenge, it will help to apply Lewisian CDT to the Chewcomb problems. Here, I turn first to conditional Chew-

these complexities accounted for, many versions of CDT will recommend that agents with certain credences bet heads in conditional Chewcomb and so the problem to be outlined below appears to remain live. I discuss the full implications of these complications shortly (primarily in discussing Rabinowiczian CDT's guidance in this case).

[25]There's a sense in which the agent does have the prophetic evidence pre-decision: Satan revelation occurs pre-decision. I return to this issue shortly but for now note simply that it's not clear cut that PROPHETIC RELEVANCE here requires a bet on tails. As such, conditional Chewcomb alone is not a clear counterexample to CDT.

comb.[26]

As with the pauper's problem, in this case the agent will have credence in just one dependency hypothesis. If we let B_H, B_T and B_N stand for betting on heads, tails or neither, respectively, and H and T stand for the chances of the coin coming up heads or tails, then this dependency hypothesis is:

K_2: $(B_H \mathbin{\square\!\!\rightarrow} H = T = 50\%) \wedge (B_T \mathbin{\square\!\!\rightarrow} H = T = 50\%) \wedge (B_N \mathbin{\square\!\!\rightarrow} H = T = 50\%)$

According to K_2, the chances of heads and tails are 50%, regardless of the pauper's decision. Now as K_2 is the only dependency hypothesis in which the agent has positive credence, Lewis's theory reduces to EDT in this case and so calculates the expected utility of each act, A, as: $\sum_W Cr(W \mid A)V(W)$. Now, after conditioning on the decision to take either bet, the agent will assign a credence of 0.99 to tails worlds (that is, worlds in which the coin comes up tails) and just 0.01 to heads worlds. Given the payouts to winning ($50 for a tails bet, $100 for a heads bet), this means that the expected utilities of the decisions to bet are:

$$EU(B_H) = (0.99 * 0) + (0.01 * 100) = 1$$
$$EU(B_T) = (0.99 * 50) + (0.01 * 0) = 49.5$$

Further, not making either bet has a guaranteed pay-off of $0 and so has an EU of 0. As such, it follows that Lewisian decision theory labels tails as rationally required in conditional Chewcomb (49.5 is higher than both 1 and 0). Further, it also labels tails as rationally required in unconditional Chewcomb for much the same reasons (49.5 is still higher than 1).

[26]My discussion here diverges from Price's §3 as Price applies only an approximately-correct version of Lewis's theory.

Lewisian CDT, then, endorses tails in both conditional and unconditional Chewcomb and, in doing so, rejects EVIDENTIAL IRRELEVANCE and follows PROPHETIC RELEVANCE. This is precisely what Price's argument suggested a theory of choice should do.

6 Problems for Lewis

The previous section established that the Chewcomb problems are not counterexamples to Lewisian CDT (which endorses tails in both cases, as desired) and so, it might seem, the promise of a general objection to CDT has already proven false. However, this is too fast. In fact, while the Chewcomb problems are not *counterexamples* to Lewisian CDT, they nevertheless pose a problem for proponents of this theory. After all, these proponents typically take the mere evidential correlations between the agent's decisions and the contents of the opaque box in Newcomb's problem to be irrelevant to rationality in this case. However, Lewisian CDT takes account of such a correlation in conditional Chewcomb and so we are owed an explanation of why this correlation is relevant in one case but not the other.

The problem for Lewisian CDT is that no principled explanation is likely to underpin the offered guidance. After all, Lewisian CDT only accounts for mere correlation in conditional Chewcomb because it takes correlation into account within single dependency hypotheses. Further, Lewis is only comfortable with this because, within dependency hypotheses and in the cases Lewis is considering (that is, normal cases), these correlations capture the causal influence of the action (Lewis, 1981, p. 11, emphasis my own):

> Within a single dependency hypothesis, so to speak, V-maximising [which takes into account correlation between decision and world]

> is right. It is rational to seek good news by doing that which, according to the dependency hypothesis you believe, most tends to produce good results. *That is the same as seeking good results.*

In abnormal cases, however, the correlation between decision and world need not capture the decision's causal influence and so deciding for these correlations is not the same as "seeking good results". That Lewis's theory accounts for these correlations, then, seems accidental, rather than principled. I take this to be the challenge raised by Price's discussion: even if Lewisian CDT gets the Chewcomb problems right, Lewis will struggle to justify the manner in which it does so without undermining the assumption that CDT gets Newcomb's problem right.

This difficulty can be made stark by considering another case (from Price, 2012, pp. 505–506), which I will call the chancy Newcomb's problem.[27] This can be outlined as follows: as in Newcomb's problem, an agent is offered the choice of either just an opaque box (one-boxing) or of both the opaque box and a transparent box containing $1000 (two-boxing). This time, however, the opaque box is filled based not on a prediction but on the toss of a fair coin: if the coin comes up heads, nothing will be placed in this box, if it comes up tails, $1,000,000 will be. Finally, Satan tells the agent that on those occasions where she two-boxes, the result of the coin toss will be heads (and hence the opaque box will be empty) and on those occasions where she one-boxes, the result will be tails (and hence the opaque box will contain $1,000,000).

This case is extremely similar to Newcomb's problem: in both cases, the agent's decision provides evidence regarding, but does not causally influence,

[27] This name does violence to Price's own naming conventions, according to which this case is Boxy Chewcomb and "Chewcomb's problem" is itself short for "chancy Newcomb's problem". But never mind (my naming system helps me avoid some ambiguities).

the opaque box's contents. More importantly, one of the key arguments for two-boxing in Newcomb's problem also applies here: (1) the agent cannot influence the contents of the box; (2) however the boxes are filled, the agent ends up better off two-boxing than one-boxing (she gets an extra $1,000); and so; (3) two-boxing is rationally required. Given that the proponent of CDT typically accepts this argument in Newcomb's problem, it seems she must also do so in the chancy Newcomb's problem and so must endorse two-boxing in this case.

A problem now arises: Lewisian CDT endorses one-boxing in the chancy Newcomb's problem. After all, in this case there is only one dependency hypothesis (as the agent is sure about the relevant chances) and so Lewisian CDT reduces to EDT here. Now, one-boxing provides the best evidence in this case (evidence that the coin came up tails and hence that the opaque box contains $1,000,000 while two-boxing provides evidence that the coin came up heads and hence that the opaque box is empty). As such, EDT (and hence Lewisian CDT) will endorse one-boxing, contra the above requirement.[28] We have made progress: both the Chewcomb problems and the chancy Newcomb's problem undermine Lewisian CDT.

7 Chewcomb, Sobel, Rabinowicz

As noted in footnote 6, Lewisian CDT is not meant to apply to abnormal cases like the Chewcomb problems nor is it the only game in CDT-town. As such, the result thus far is hardly deeply problematic for CDT. For the Chewcomb problems to be of deep interest, they will need to pose a problem

[28]The version of CDT in Hitchcock, 2015 also endorses one-boxing here and comes paired with an argument for why this is appropriate. I won't evaluate this argument here but note that this same argument entails the rationality of one-boxing in cases where Lewisian CDT endorses two-boxing and so Lewisian CDT faces problems either way.

not just for Lewisian CDT but for other versions of CDT too (in the remainder of the paper I set aside the chancy Newcomb's problem as it poses no problem for either Sobelian or Rabinowiczian CDT). I turn, then, to other versions of CDT.

Start with Sobelian CDT. Unconditional Chewcomb alone is enough to undermine this theory. After all, Sobelian CDT follows the chances and so wipes out the prophetic evidence in this case. As such, in calculating the EU of each decision, it will assign a probability of 0.5 to the coin coming up heads and 0.5 to it coming up tails and so endorse betting heads (as Sobelian CDT sees both bets as equally likely to pay out and a heads bet pays twice as much as a tails bet). However, PROPHETIC RELEVANCE entails the rational requirement of choosing tails and so Sobelian CDT is problematic.[29] The Chewcomb problems, then, mediate against Sobelian, as well as Lewisian, CDT. So far so good (for Price's argument).

However, now consider Rabinowiczian CDT and conditional Chewcomb (which is the Chewcomb case most likely to cause problems for this theory):[30] in this case, Rabinowiczian CDT entails that what is rational depends on how the agent believes she will act. In particular, if the agent is sure that she will bet heads then, given that $Cr(T \mid B_H) = 0.99$ (that is, given Satan's information), she will assign 99% of her credence to tails worlds and 1% to heads worlds. Further, on the Rabinowiczian approach to CDT, imaging on any of the decisions will not change this proportion (because the outcome of the chance process is independent of the agent's decision). Given that these

[29]This is too quick. PROPHETIC RELEVANCE, a principle of epistemic rationality, merely requires that the agent set her credences a certain way while Sobelian CDT, as a principle of practical rationality, merely tells an agent how to decide given her credences (and does not restrict how these credences are set). Nevertheless, I think it clear that heads is irrational in unconditional Chewcomb and so clear that Sobelian CDT is problematic.

[30]Rabinowiczian CDT endorses tails in unconditional Chewcomb and so no problem arises here.

credences are the same as the credences relied upon by Lewisian CDT and given that the values of worlds will also be the same as in this discussion, this means that Rabinowiczian CDT provides the same guidance as Lewisian CDT: it endorses betting tails and so acts as Price suggests (at least if the agent is certain that she will bet heads: I consider other cases shortly).[31]

Rabinowiczian CDT, however, has one advantage and one disadvantage over the Lewisian account (which also endorsed tails). First, the advantage. This theory doesn't take the evidence provided by the decision to be directly relevant to rationality in the Chewcomb problems but rather takes it, along with the agent's beliefs about how she will decide, to determine the agent's pre-decision credences. These pre-decision credences are then taken to be directly relevant to rationality. Given that the position adopted by the proponent of CDT with regards to Newcomb's problem in no way requires her to demand that agents fail to update their credences in this manner, this means that there is no inconsistency between Rabinowiczian CDT's guidance in the Chewcomb problems and in Newcomb's problem (and hence means that Rabinowiczian CDT can be justified in a way that Lewisian CDT could not). To put this another way: unlike Lewisian CDT, Rabinowiczian CDT can perform appropriately in the Chewcomb problems without violating EVIDENTIAL IRRELEVANCE (because it uses the evidence that would be provided by the decision only to determine the agent's unconditional credences and doesn't, in addition, take this evidence to have further relevance once these credences have been set: see footnote 20).[32]

[31]Some deny that a deliberating agent can have credences regarding her decision (cf. Price, 2012, pp. 528–531) and Rabinowiczian CDT may endorse heads if so. I've never found this claim regarding deliberation particularly plausible but Rabinowiczian CDT weathers the storm regardless: see the first response to the disadvantage discussed below.

[32]So this response denies that it's possible for Satan's revelation to influence the agent's conditional credences without also influencing her unconditional credences (and as CDT responds to these unconditional credences it can then account for Satan's information).

Now the disadvantage: Rabinowiczian CDT doesn't always entail the rationality of betting tails in conditional Chewcomb. In particular, in some cases where the agent has positive credence that she will not take either bet, this theory will label betting heads as rational. We can see this in the extreme case where the agent is sure that she'll refuse both bets. In this case she ignores Satan's information (as she considers it irrelevant to her) and so assigns 50% of her credence to heads worlds and 50% to tails worlds (and this proportion remains the same after imaging). Given these beliefs, Rabinowiczian CDT labels betting heads as rational. After all, according to these beliefs the coin has the same chance of coming up either heads or tails and yet a successful heads bet pays twice as much as a successful tails bet. So while Rabinowiczian CDT always labels betting tails as rational in unconditional Chewcomb, it sometimes labels betting heads as rational in conditional Chewcomb. Price's problem reasserts itself: surely the same decisions are rational in both cases so this theory must be flawed.

A first response: the equivalence claim (according to which the same decisions must be rational in the two cases) can be challenged. Decision theory is a theory of subjective rationality and what is subjectively rational depends on an agent's beliefs. However, in conditional Chewcomb, the addition of the no bet option opens up the possibility of the agent being in epistemic positions that won't arise in unconditional Chewcomb and so when such a situation arises there's no reason to think that the same decision must be rational in both cases (because the agent will be in different epistemic positions in the two cases and different things may be subjectively rational from different epistemic positions). More particularly, in unconditional Chewcomb the agent's evidence entails that the coin will almost certainly come up tails but in conditional Chewcomb this need not be the case (it

won't be if the agent thinks she's unlikely to take either bet). These are the circumstances under which Rabinowiczian CDT endorses different decisions in the two cases and in these circumstances there are grounds to think that different decisions may be rational.[33]

A further response: if we accept the view of CDT defended in Joyce, 2012, Rabinowiczian CDT will in fact never endorse betting heads in conditional Chewcomb.[34] In particular, Joyce argues that CDT involves two claims: (1) a claim that at any time, t, options should be evaluated according to their causal expected utility; and (2) a claim that an agent should make a decision in accordance with an evaluation at a time t if and only if her evaluation has taken into account all of the relevant information that is readily available at t. We can apply this to the version of conditional Chewcomb where the agent starts out confident that she won't bet. According to Rabinowiczian CDT, betting heads initially has a higher EU than betting tails in this case. However, on the Joycean view, this provides the agent with evidence that she will bet heads and this evidence hasn't been accounted for so the agent shouldn't decide in accordance with this evaluation.[35] Instead, she should take into account this evidence and then re-evaluate the expected utility of the options. Having done so, she will calculate the EUs in light of the

[33]The contrary intuition plausibly follows from the undesirability of refusing both bets and the assumption that the addition of an inferior option should not change what is rational. But this misses the mark: the decision's influence results not from its desirability (or lack thereof) but from its impact on the agent's epistemic position. There's no reason to think the decision's undesirability makes its epistemic impact irrelevant. A similar response can also be given to attempts to bolster this intuition via an independence principle, as in Price's footnote 11 (cf. Luce and Raiffa, 1967, p. 288 to see why the required independence principle is implausible).

[34]See also Arntzenius, 2008 and Skyrms, 1990.

[35]Of course, this is only *clearly* true if the agent is a (causal) expected utility maximiser (otherwise the EU calculations need not change her credences regarding her future behaviour). Still, the argument can be extended to many other cases and where this is not possible the first outlined defence of Rabinowiczian CDT is likely to suffice (as these cases will typically involve irrational agents or agents in impoverished epistemic states: in such cases, this defence is particularly compelling).

assumption that she is likely to bet heads and so tails will end up with the highest EU. Again, then, the agent has information that she has not taken into account. She should now take into account this evidence and re-evaluate the options. She will do so in light of the assumption that she will bet tails and so will think the coin is likely to come up tails. As such, tails will get the highest EU. The process now stops as she has already taken into account the evidence provided by tails having the highest EU.[36] As such, the agent should now decide in accordance with this evaluation and so should bet tails. Even if the agent starts out thinking she will refuse both bets, then, Rabinowiczian CDT interpreted in the Joycean manner will entail the rationality of betting tails and so will avoid Price's challenge.[37]

It follows that Rabinowiczian CDT gets things right in the Chewcomb problems and so these cases do not undermine CDT generally but rather undermine Lewisian and Sobelian CDT in particular (at least as currently construed: see footnote 39).

[36]In fact, while the agent has taken into account that tails has the highest EU, she may not have taken into account tails' exact EU (which may have changed since the previous evaluation). More accurately, then: when this process terminates depends on the method used to update credences based on EU evaluations (an issue that is too large to resolve here). Nevertheless, given plausible assumptions about this method, there will be an equilibrium point at which no new information is provided by the latest evaluation (cf. Arntzenius, 2008, pp. 293–294). Further, to be reasonable an update method must: (a) reach equilibrium; and (b) entail that tails has the unique highest EU at equilibrium (after all, tails is the only decision that is rational in light of the assumption that this decision will be made).

[37]Two comments. First, this modification of Rabinowiczian CDT will not change the guidance that this theory provides in the pauper's problem (as here the fact that some decision receives the highest EU in the initial evaluation merely reinforces the rationality of this decision). Second, in discussion Brian Weatherson has noted that this view's reliance on equilibrium within an agents belief might be problematic given that in game-theoretic contexts some equilibria are intuitively irrational (cf. Cho and Kreps, 1987). A full evaluation of this issue is beyond the purview of this paper (though I think CDT may well have the tools to resolve this issue) so here I simply note that even if equilibrium-style reasoning is problematic the first defence of Rabinowiczian CDT suffices to defend the theory.

8 Conclusions

At this point the forest is in danger of being lost for the trees. A reminder, then, of the territory covered: I have discussed four abnormal cases (the pauper's problem, the two Chewcomb problems and the chancy Newcomb's problem) and argued for two conclusions. First, I have argued that neither the pauper's problem nor Price's cases undermine CDT generally, at least when each set of cases is considering in isolation.[38] Second, I have argued that these cases do nevertheless mediate between different versions of this theory, pushing us toward Rabinowiczian CDT (or, at least, toward some member of a class of theories that contains, at minimum, Rabinowiczian CDT).[39] I conclude that while Lewis is right that these cases are problematic for CDT, they are not irresolvably problematic: CDT survives the presented cases.

[38] I have not considered whether the cases *jointly* undermine CDT: Price's cases push us to Rabinowiczian CDT so if this theory gets the pauper's problem wrong then the cases might, in combination, pose a general problem for CDT. Four responses: (1) the Rabinowiczian response to the pauper's problem seems plausible to me so I remain unconcerned; (2) if CDT is unproblematic in both Newcomb's problem *and* the cases in (Egan, 2007) then the Rabinowiczian line seems even more plausible (and if not, CDT is independently flawed anyway); (3) I think the pauper's problem complex enough that I am loathe to use it to mediate between theories of choice so question the strength of this argument; and (4) given the radical divergence of versions of CDT in abnormal cases, it seems likely that there will exist versions that side with each plausible view on the pauper's problem while reasoning appropriately in Price's cases (for example, perhaps the version of CDT in Hitchcock, 2015 can play a role in this project). If so the challenge disappears.

[39] Other potential members of this class: forms of interventionist CDT (Hitchcock, 2015) and modified Sobelian and Lewisian CDT that either take chances to be strange in cases involving prophetic evidence or directly factor prophetic evidence into their calculations.

References

Arntzenius, Frank. 2008. "No Regrets, or: Edith Piaf Revamps Decision Theory". *Erkenntnis* 68: 277–297.

Cho, In-Koo, and David M. Kreps. 1987. "Signaling Games and Stable Equilibria". *The Quarterly Journal of Economics* 102: pp. 179–222.

Egan, Andy. 2007. "Some Counterexamples to Causal Decision Theory". *The Philosophical Review* 116: pp. 93–114.

Hitchcock, Christopher. 2015. "Conditioning, intervening, and decision". *Synthese* : 1–20.

Joyce, James M. 2012. "Regret and Instability in Causal Decision Theory". *Synthese* 187: 123–145.

Lewis, David. 1979. "Counterfactual Dependence and Time's Arrow". *Noûs* 13: 455–476.

———. 1980. "A Subjectivist's Guide to Objective Chance". In *Studies in Inductive Logic and Probability*, edited by Richard C. Jeffrey, vol. 2. University of California Press, 83–132.

———. 1981. "Causal Decision Theory". *Australasian Journal of Philosophy* 59: 5–30.

———. 1982. "Letter to Wlodek Rabinowicz".

Luce, R.D., and H. Raiffa. 1967. *Games and Decisions: Introduction and Critical Survey; a Study of the Behavioral Models Project*. Wiley.

Nozick, Robert. 1969. "Newcomb's Problem and Two Principles of Choice".

In *Essays in Honor of Carl G. Hempel*, edited by Nicholas Rescher. Reidel, 114–146.

Price, Huw. 2012. "Causation, Chance, and the Rational Significance of Supernatural Evidence". *Philosophical Review* 121: 483–538.

Rabinowicz, Wlodek. 1982a. "Letter to David Lewis".

———. 2009. "Letters From Long Ago: On Causal Decision Theory and Centered Chances". In *Logic, Ethics, and All That Jazz - Essays in Honour of Jordan Howard Sobel*, edited by Johansson Lars-Göran, vol. 56. 247–273.

Rabinowicz, Wlodzimierz. 1982b. "Two Causal Decision Theories: Lewis vs Sobel". In *Philosophical Essays Dedicated to Lennart Åqvist*, edited by Tom Pauli. Philosophical Society and Dept. Of Philosophy, University of Uppsala, 299–321.

Skyrms, Brian. 1990. *The Dynamics of Rational Deliberation*. Harvard University Press.

Sobel, Jordan Howard. 1980. *Probability, Chance and Choice*. Unpublished.