Can Free Evidence Be Bad?: Value of Information for the Imprecise Probabilist

Seamus Bradley and Katie Steele

September 8, 2015

Abstract

This paper considers a puzzling conflict between two positions that are each compelling: (A) it is irrational for an agent to pay to avoid 'free' evidence, and (B) rational agents may have imprecise beliefs. An important aspect of generating this conflict is resolving the question of how rational (imprecise) agents ought to make sequences of decisions – we make explicit what the key alternatives are and defend our own approach. We endorse a resolution of the aforementioned puzzle – we privilege decision theories that merely permit avoiding free evidence over decision theories which make avoiding free information obligatory.

1 Introduction

If evidence is available for 'free'[1], it is presumably a good idea to pursue that evidence and take it into account when making decisions. Good (1967) proves as much for the case where Your[2] degrees of belief are precise. That is, Good proves that when Your beliefs or credence can be represented by a single probability function, free evidence cannot be detrimental. Good proves that, under standard conditions, Your expectation increases (or at least cannot decrease) if You pursue free evidence.[3]

In the case where Your credence is imprecise, however, no such nice result exists. Indeed, several authors have shown that, in such cases, You may effectively pay to avoid free evidence (Kadane, Schervish, and Seidenfeld 2008; Grünwald and Halpern 2004). These examples rely on the phenomenon of dilation (Seidenfeld and Wasserman 1993; Pedersen and Wheeler 2014). What happens in dilation is that conditionalising on some evidence can cause an event whose prior probability was sharp to have an unsharp posterior probability. This 'fuzzying' of Your degrees of belief may lead to Your paying to avoid free evidence. This paper investigates the *puzzle* suggested by this result.

Let us quickly set the scene. Many have argued, convincingly, that rationality permits imprecision: a rational agent may have incomplete preferences that are best represented by a set of probability- and utility-function pairs, rather than a single, precise such pair.[4] On the other hand, rationality surely requires that an agent *not* pay to avoid free evidence. If all plausible decision theories for handling imprecision have the consequence that You may pay to avoid free evidence, we have a contradiction, and must consider which of the

[1] What exactly it means for evidence to be free will be discussed later.

[2] 'You' is the arbitrary intentional agent who is under discussion. This practice begins, appropriately, with Good.

[3] A posthumous publication of Ramsey's work contains a precursor of the theorem (Ramsey 1990). In the preamble to the publication, Sahlin also notes a precursor in Savage (1972 [1954]). Moreover, Kadane, Schervish, and Seidenfeld (2008) cite a version of the theorem in Raiffa and Schlaifer (1961).

[4] Formal treatments of *imprecise probabilism* can be found in Walley (1991) and Augustin et al. (2014). Philosophical defenses of the view can be found in Levi (1974), Joyce (2005) and Sturgeon (2008). Note that this paper focuses on imprecise beliefs, and we assume utilities of basic outcomes are precise. Moreover, we restrict attention to imprecision as represented by sets of probabilities, but we do not deny that there are alternative representations of imprecision/uncertainty, such as Dempster-Shafer functions or upper and lower probabilities. For a thorough analysis of the different representations of uncertainty, see Halpern (2003).

premises should be given up.

The next section presents the apparent *trilemma* in more detail: we initially outline the set-up for, and statement of, Good's theorem, and show, by example and in line with other authors, that his lesson about the invariable goodness of free evidence does not straightforwardly extend to the imprecise context. While not an original aspect of the paper, we think it important to carefully explain Good's reasoning about free evidence, and its prima facie extension to the imprecise realm, because this is a noteworthy and interesting issue for decision theory that deserves attention. We go on to clarify the premises of our proposed trilemma. This serves to highlight the main issue of the paper: whether *all* plausible decision theories for handling imprecision do in fact fall afoul of free-evidence intuitions.

In order to discuss the commitments of different decision rules regarding free evidence, it is necessary to first elaborate, in Section 3, how You should negotiate a dynamic- or sequential-decision problem, i.e. a decision problem that involves choices both now and in the future. We do not dispute the orthodox *sophisticated-choice* approach, but we show that it has not been adequately spelled out in the current literature, and the missing details leave room for controversy that is relevant to our free-evidence trilemma. So we effectively introduce and argue for our own version of sophisticated choice for the imprecise probabilist. In a sense, this middle section is the most important contribution of the paper.

Section 4 then examines an important, perhaps even the only plausible, decision rule for handling imprecision; this is the maximally permissive non-dominated-set (NDS) rule. The NDS rule does not preclude paying to avoid free evidence, but we prove that it at least never mandates such payment. Section 5 comments on how this result fits into the broader decision theory literature on dynamic coherence. In particular, we note that previous work in the literature shows that no decision rule can do better than the NDS rule with respect to the pursuit of free evidence. Section 6 summarises our stance on the trilemma.

2 Good's theorem and troubles for the imprecise probabilist

I.J. Good showed that it always pays in expectation for the precise probabilist to pursue free evidence. The proof can be found in Good (1967). A key assumption is that You conform to standard Bayesian decision theory, i.e., You are an expected utility maximiser and update Your beliefs in line with

the rule of conditionalisation.[5] Good's result accords with intuitions about seeking evidence and experimentation being advantageous. Indeed David Miller (1994) notes that one of A.J. Ayer's two main criteria for a successful account of scientific method is that the account should advise scientists to pursue free evidence. (The second criterion is that the account should advise the scientist to base inferences on all the evidence available.) Bayesian decision theory thus meets Ayer's standards because it is not inconsistent with the two criteria (notwithstanding Miller's arguments to the contrary). Carnap also discussed what he called the 'Principle of Total Evidence' which amounted to the claim that reasoners should not ignore available evidence when estimating a probability.[6]

The problem set-up for the proof is as follows: There is a partition of the state space $H_1, \ldots, H_r$ and some acts $A_1, \ldots, A_s$ which can have different payouts for the different events H_i. Your utility for an act in a particular event is: $U(A_j(H_i))$. Your credences over the events are $\Pr(H_i)$. There is another partition of the state space $E_1, \ldots, E_t$. You are offered the chance to learn which E_k obtains. Pursuing the new evidence has no disutility associated with it so the utilities of the basic outcomes in the decision problem are the same whether or not You learn. This is essentially the meaning of free evidence.[7] Good shows that, under these assumptions, opting for learning which E_k obtains has at least as high expectation as not learning, and learning always has higher expectation when it is possible that the new evidence may change Your choice of act amongst $A_1, \ldots, A_s$.

The idea behind the proof is illustrated in Table 1.[8] Not learning effectively

[5]The requirement that updating be by conditionalisation can be generalised: Skyrms (1990) and Huttegger (2013).

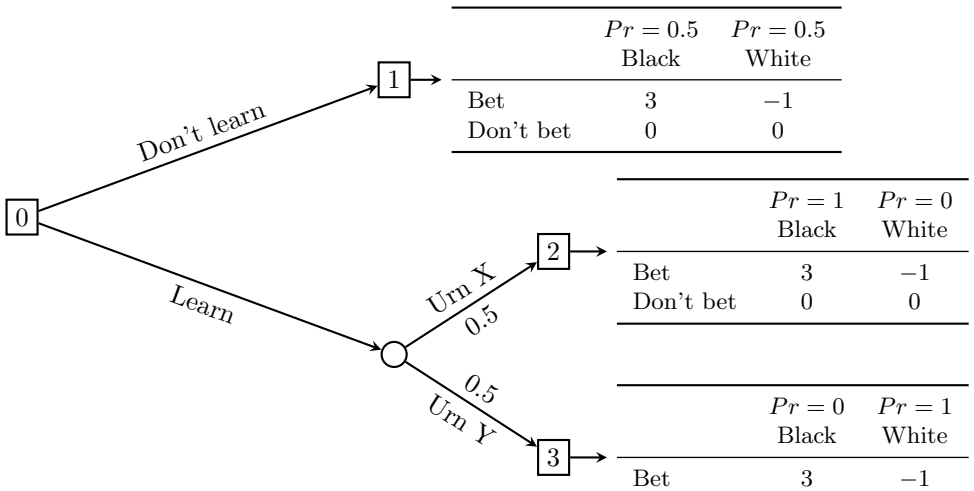
[6]Good in fact claimed that his theorem shows that Carnap's 'Principle of Total Evidence' (PTE) is a consequence of Bayesian decision theory. We think this confuses the maxim that 'one should always base one's beliefs on the totality of one's evidence' with the maxim that 'one should always seek new potential evidence if it is cost-free'. The former maxim, the PTE, concerns the set-up of a Bayesian model, or else it is simply an aspect of probabilism (cf. Miller (1994), p. 159). By contrast, the second maxim is a consequence of Bayesian decision theory.

[7]Note that this condition excludes salient cases where apparently free evidence is not expected to be beneficial, such as Your friend offering to tell You the end of the novel You are reading, or the details of Your surprise party. In these cases, learning is not in fact free, if it is properly modelled, because the knowledge changes the utility of the outcomes directly.

[8]Strictly speaking, each cell of the columns labelled E_1 or E_t should look like this: $\sum \Pr(H_i|E_1)U(A_1(H_i \cap E_1))$ since those are the basic elements of the space. This would have made the table much more unwieldy, and not added to the exposition. In short, we have made the assumption the utilities of the consequences of each act are not dependent

	E_1	...	E_t	Average
A_1	$\sum P(H_i\|E_1)U(A_1(H_i))$	...	$\sum P(H_i\|E_t)U(A_1(H_i))$	$\sum P(H_i)U(A_1(H_i))$
A_2	$\sum P(H_i\|E_1)U(A_2(H_i))$	...	$\sum P(H_i\|E_t)U(A_2(H_i))$	$\sum P(H_i)U(A_2(H_i))$
$\vdots$		$\ddots$		$\vdots$
A_s	$\sum P(H_i\|E_1)U(A_s(H_i))$	...	$\sum P(H_i\|E_t)U(A_s(H_i))$	$\sum P(H_i)U(A_s(H_i))$

Table 1: The idea behind the proof



	$Pr = 0.5$ Black	$Pr = 0.5$ White
Bet	3	−1
Don't bet	0	0

	$Pr = 1$ Black	$Pr = 0$ White
Bet	3	−1
Don't bet	0	0

	$Pr = 0$ Black	$Pr = 1$ White
Bet	3	−1
Don't bet	0	0

Figure 1: A simple decision problem

means choosing among the averages of the acts conditional on each E_k. Your expectation is thus the maximum value in the final 'Average' column. Your expectation for learning, on the other hand, is the average of the maximums in each of the columns $E_1, \ldots, E_t$. In short, learning is always at least as good as not learning because the average of maximums is always at least as high as the maximum average.

Consider the following simple example, depicted in figure 1. Note that the squares indicate *choice nodes*: places where You must make a choice; and the circles indicate *chance nodes*: uncertainty that is then resolved one way or another.[9] There are two urns labelled X and Y. You believe that urn X contains 10 black marbles and that urn Y contains 10 white marbles. One urn will be selected at random by the toss of a fair coin and a marble drawn

on which E_k occurs, given H_i. Nothing important hangs on this simplification of the example.

[9]One can think of these circular *chance nodes* as 'Nature's choice nodes'.

from it. You are offered a 2 to 1 bet on black: if black is drawn You end up with 3 and if white is drawn You lose your stake of 1. It should be obvious that learning (i.e. 'down' at choice point 0) has higher expectation in this case: if You were to learn that the draw is from X, You would bet on black and predict a win of 3, and if You were to learn that the draw is from Y, You would refrain from betting and receive 0. Thus Your expectation would be $0.5 \times 3 + 0.5 \times 0 = 1.5$, which is greater than 1, the expected utility of not learning (since if You chose up at node 0 You would choose to bet).

It is worth noting that Good's proof concerns the *expected* value of free evidence. It may be the case that You 'get unlucky', and Your expected utility in fact goes down upon receiving some evidence. For instance, You may learn that the draw is from Y in our example, giving an expectation of 0 (since if You learned Y You would choose not to bet), which is of course less than 1. This is a possible eventuality; nonetheless, the *expected* utility of learning free evidence (i.e. before a particular piece of evidence is actually received) is greater than or equal to not learning.

Consider a slightly different example: You believe that urn X definitely contains 8 black and 2 white marbles, and that urn Y contains 2 black and 8 white. Again, learning which urn is drawn from has higher expectation, as You would tailor Your choice of 'Bet'/'Don't bet' to the new evidence, as before. Here You would not be guaranteed the prize, by Your own lights, even if You learn that the draw is from urn X, but Your choice will accord with what You take to be the proportions of balls in the urn, whatever urn is revealed, and so overall You have a higher expectation.

There is a problem when we try to apply this to the imprecise case. That is, where Your belief is represented by a *set* of probability measures, $\mathcal{P}$, rather than just one measure. Call this set Your *representor*. Imagine a scenario similar to the last one. Imagine You believe that there are a total of 10 black and 10 white marbles distributed somehow among the urns X and Y. Each urn contains 10 marbles. An urn will be selected at random by flipping a fair coin, and a marble drawn from it. Using X and Y to refer to the propositions 'the marble is drawn from urn X' and 'the marble is drawn from urn Y', respectively; and using B and W to stand for the propositions 'the marble drawn is black' and 'the marble drawn is white', respectively; in the spirit of imprecise probabilism, the following is a plausible characteristic of Your belief representor, $\mathcal{P}$: $\Pr(B|X) = 1 - \Pr(B|Y)$ for all $\Pr \in \mathcal{P}$. As such, Your credences before learning regarding the colour of the marble drawn are as follows: $\mathcal{P}(W) = \{0.5\} = \mathcal{P}(B)$. You believe that the number of white marbles and the number of black marbles are

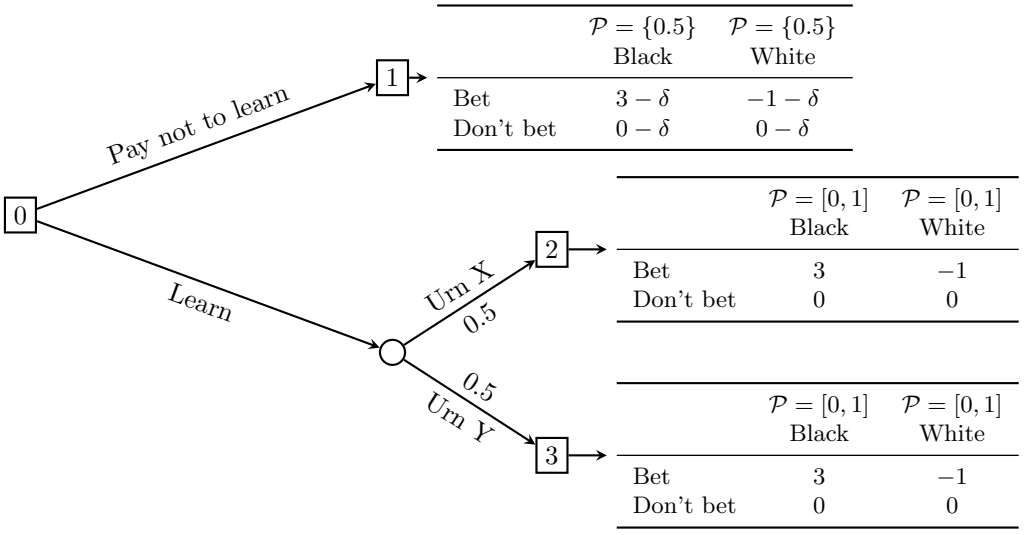
equal and that over the two urns their probabilities average out. It seems plausible that Your conditional credences are, however, imprecise: You have no information about how the marbles are distributed between the urns, and so Your representor, $\mathcal{P}$, plausibly includes the possibility that urn X certainly contains only white marbles and also the possibility that X certainly contains no white marbles, and everything in between. That is, $\mathcal{P}(W|X) = \mathcal{P}(B|X) = \mathcal{P}(W|Y) = \mathcal{P}(B|Y) = \{0, \frac{1}{10}, \ldots, \frac{9}{10}, 1\}$, or, if *convexivity* were mandated (as per Levi (1974) and Levi (1980)), Your representor would plausibly have $\mathcal{P}(W|X) = [0, 1]$, and likewise for the other conditional attitudes.[10] For reasons of notational convenience we will stick to using $[0, 1]$ to represent the aforesaid imprecise conditional beliefs. So learning which urn is drawn from *dilates* Your probability for white from 0.5 to $[0, 1]$, and likewise for black. Note that this analysis of the problem assumes the standard rule for updating imprecise beliefs, known as *generalised conditioning*, which basically amounts to pointwise conditioning: After learning X, say, $\{\Pr(-|X) : \Pr \in \mathcal{P}\}$ is Your new representor.[11]

Recall that You are offered a 2 to 1 bet on black: if black is drawn You win 3 and if white is drawn You lose 1. Before learning, it is clear that this bet is advantageous and You should take the bet, rather than stick with the status quo. (As before, Your expectation for not learning, given the credences described above, is 1.) If You choose to learn, on the other hand, Your expectation for the bet will inevitably dilate from $\{1\}$ to $[-1, 3]$. Whether or not the revealed urn is X or Y, the bet no longer has higher than or equivalent expected utility to the status quo: the two acts in each case become incommensurable. Pursuing free evidence is no longer straightforwardly better for You, even though it may lead to You changing Your choice of option.

Now imagine the same problem, except that choosing not to learn comes with a small penalty of δ (where $\delta > 0$). This makes the problem more vivid, because now, if You choose not to learn, You are in fact *paying* to avoid learning. Figure 2 depicts this revised problem. You face a choice at the initial node (labelled 0) between learning and paying not to learn. let's assume Your credences are as before. Learning leads to evidence about the coin flip – whether the urn to be drawn from is X or Y. As per Good's theorem, the final decision problem is constant, whatever happens, except that the option of not learning involves δ utility subtracted from all outcomes.

[10] Whether convexity *is* mandated when in fact the objective probability of drawing white given X could not possibly be, say, $\frac{6}{21}$, is a tricky question and one we shall ignore. Nothing in our discussion hinges on the sets of probabilities being convex.

[11] Refer to Bradley and Steele (2014b) for further discussion of this example, in particular whether dilation is peculiar from a purely epistemic point of view.



	$\mathcal{P} = \{0.5\}$ Black	$\mathcal{P} = \{0.5\}$ White
Bet	$3 - \delta$	$-1 - \delta$
Don't bet	$0 - \delta$	$0 - \delta$

	$\mathcal{P} = [0, 1]$ Black	$\mathcal{P} = [0, 1]$ White
Bet	3	-1
Don't bet	0	0

	$\mathcal{P} = [0, 1]$ Black	$\mathcal{P} = [0, 1]$ White
Bet	3	-1
Don't bet	0	0

Figure 2: The problematic decision problem

If You choose not to learn, You effectively pay δ to avoid free evidence.[12]

Imagine, for instance, that You conform to the so-called *gamma-maximin* rule, which holds that one should always choose the option that has the largest minimum expected utility.[13] You would reason as follows: if learning is chosen, there will inevitably be a choice between betting, with expected utility $[-1, 3]$, and not betting, with expected utility 0. The latter has the greater minimum expected utility, so learning effectively leads to not taking the bet, which has 0 utility. (This process of working backwards through a decision problem is referred to as *sophisticated* reasoning or backward induction. All of the examples here have involved sophisticated reasoning. We say more about this in the next section.) Not learning, on the other hand, will mean that the bet is chosen, and this has expected utility $1 - \delta$ from the initial vantage point. Assuming $\delta < 1$, not learning is therefore preferable. You thus pay to avoid free evidence.

By way of setting the stage for the remainder of the paper, let us be explicit about the proposed trilemma:

[12]This is really a case of better-than-free evidence. Strictly speaking, learning *does* change the utilities of the outcomes, so this is not really *free* evidence, as defined earlier. However, the changes are uniformly for the better, so it's still the case that You do not pay to acquire the evidence. We call taking the 'not learning' branch 'paying to avoid free evidence' when we should really say 'choosing not to learn better-than-free evidence'.

[13]Gärdenfors and Sahlin (1982), for instance, defend this rule. Seidenfeld (2004) refers to the rule as gamma-maximin, and we follow this terminology here.

$P1$ *Paying to avoid free evidence is irrational.*

$P2$ *Incomplete preferences are not irrational, and can be represented by sets of probabilities and utilities. In other words, imprecise probabilism is permissible.*

$P3$ *All plausible decision theories for handling imprecision sanction paying to avoid free evidence.*

As mentioned earlier, $P2$ is defended elsewhere. We will revisit $P1$ later, but let us simply note, for now, that it is hard to conceive how 'money' spent on avoiding free evidence could be money well spent. Decision theory aside, any sensible person would surely ignore new free evidence that they might obtain, if it is really unhelpful, rather than pay to remain ignorant about it.

The initial focus of this section is the final premise, $P3$. What we have illustrated above is just that a very specific type of imprecise probabilist – one whose belief-updating rule is *generalised conditioning* and who chooses according to gamma-maximin – chooses making a payment over receiving free evidence. It remains to be seen, however, whether this is also a consequence of other decision theories for handling imprecision. Indeed, we now investigate whether there may be alternative generalisations of standard Bayesian decision theory that do not recommend this unintuitive course of action.

We will focus exclusively on alternative decision rules. That is, we do not here consider alternative rules for updating belief, but rather assume *generalised conditioning* throughout. As mentioned above, this is arguably the most natural generalisation of Bayesian conditioning for the imprecise context. Stated in full, the rule holds that Your posterior belief in some proposition H, after learning the proposition E, is just the set of conditional probabilities pertaining to all probability functions in Your representor for which the conditional is defined:

$$\mathcal{P}(H|E) = \{\Pr(H|E), \Pr \in \mathcal{P}, \Pr(E) \neq 0\} \tag{1}$$

This is the rule we implicitly assumed for the example above, and it is subject to dilation. Of course, one might suggest that, given that dilation seems to be directly implicated in paying to avoid free evidence, we should explore alternative rules for updating imprecise beliefs that might preclude dilation, and its attendant problems. Elsewhere we do explore this possibility,[14] but this is not our focus in this paper. Note just that we do not regard this to be a fruitful avenue for developing imprecise probabilism; dilation is not a good reason for departing from generalised conditioning.

[14]See Bradley and Steele (2014b)

So, as mentioned, we restrict our attention to decision rules, and hold fixed generalised conditioning as the belief-update rule. The gamma-maximin rule assumed above is already criticised on other grounds (Seidenfeld 2004), so there is hope that the bad results concerning dilation are specific to this rule. In what follows, our discussion of decision rules proceeds under the assumption that there is already an appropriate representation of Your beliefs and desires. In particular, here we assume that You have imprecise probabilistic beliefs, i.e. Your belief representor is a set of probability distributions, $\mathcal{P}$, and yet, for simplicity, You have a precise utility function over basic outcomes.[15] The claims of this paper pertain to this assumed set-up – imprecise probabilities yet precise utility function – but we do not anticipate any problems in extending the results to the case of both imprecise probabilities and utilities. Before getting to the specifics of the various decision rules, however, it is important to first specify how, in general, one should negotiate a sequential-choice problem, since the question of whether or not to pursue free evidence is a problem of this kind.

3 Sophisticated choice for the imprecise probabilist

Recall that the problem we set up in the previous section involved several ingredients: first, a particular update rule; second, a particular decision rule; and third, a particular approach to sequential choice. In the example we had: generalised conditioning, gamma-maximin, and sophisticated choice. We take the first and third of these for granted throughout. We will see, in a moment, an alternative to gamma-maximin that fares better in the kind of problems we are looking at. But first, we want to spell out a little more carefully what sophisticated choice amounts to in the current setting. For precise probabilities, sophisticated choice is the orthodoxy,[16] but for imprecise probabilities, it has not been fully fleshed out what sophistication means. We take the following discussion to flesh out the correct version of sophisticated choice.

[15] One way to conceive of the representation is as follows (cf. Levi (1986) and Kaplan (1996)): You have *incomplete preferences* which can be identified with a set of complete preference orderings that are each possible expected-utility extensions of the incomplete preferences. These extended orderings each correspond to a precise probability and utility representation. Identifying a partial order with a set of probability–utility pairs is not straightforward but it can be done: see for example Seidenfeld, Schervish, and Kadane (1995). Overall, You are thus represented by the set of these probability-utility pairs. On the basis of Your probability-utility pairs, the various decision rules specify which acts among those available are choice-worthy.

[16] Advocates of the sophisticated approach to sequential choice include Seidenfeld (1988), Levi (1991) and Maher (1992).

3.1 Sophisticated choice, user-friendly version

Recall the sequential-choice reasoning in the previous section. You look at the terminal nodes and work out which act maximises minimum expectation with respect to Your representor. This done, You now treat those terminal nodes as if the choice were fixed. From the earlier node, You treat the earlier choices as if they were choices that determinately led to the gamma-maximin-best act. You effectively act as if You know Your future self will choose that way.

Let's recast the imprecise decision problem we discussed earlier as a group decision problem where there is a common utility for the agents in the group.[17] Let's also state, in accordance with the suggested decision rule, that the group chooses by evaluating an act by the lowest expected value assigned to it by some committee member. Recall figure 2. Each agent knows that if the group ends up at node 1, the group choice will be to bet since every member agrees that the expected value of betting at node 1 is $1 - \delta$. Each agent (at node 0) also knows that if the group chooses to learn and ends up at node 2, the group will choose not to learn, since there is a committee member who assigns expected value -1 to betting: the agent who thinks all the black marbles are in urn Y (call this agent Mr. White). Likewise, at node 3 the group will choose not to bet since there is an agent who evaluates betting at -1: the agent who thinks all the black marbles are in urn X (call this agent Mrs. Black). So at node 0, any agent evaluates choosing to learn as a fifty-fifty gamble between two refusals to bet, resulting in expected value of 0. Thus every committee member agrees at node 0 that paying not to learn is better, given what they believe about how the committee will choose at later nodes. This is the problem of free evidence for the imprecise probabilist recast as a problem for group belief.

But consider the following alternative reasoning. Note that each committee member is a precise probabilist, and as such, Good's theorem applies: each committee member thinks that learning increases expectation. So each committee member agrees that choosing to learn is uniquely permissible. Since any plausible decision rule should respect unanimity of this kind, there is no problem of free evidence for the imprecise probabilist.

What has gone wrong? We have two ways of reasoning about the decision problem that You or Your credal committee face, and they seem to give

[17] Levi (1986) and Levi (1999) has argued that imprecise probabilities are an appropriate model of group belief, and that individual agents can be conflicted in the same way that groups can. See also Seidenfeld, Kadane, and Schervish (1989).

conflicting conclusions. To see what goes wrong with the latter kind of reasoning, we need to take a detour into backward induction, through Greek myths, game theory and group decision making.

On his way back from Troy, Odysseus' ship must sail past the sirens. Circe has warned Odysseus that the sirens' song will drive him and his crew mad and cause them to crash their ship on the treacherous rocks thereabouts. Wily Odysseus orders his crew to stop their ears with beeswax, Odysseus, however, is curious to hear what the sirens' song sounds like. Odysseus has the option of ordering his crew to tie him to the mast or to be left unhindered. Possible courses of action once within range of the sirens are for Odysseus to stay onboard or to dive into the sea and drown. Odysseus would rather not be tied to the mast, but prefers that to drowning. He knows that the sirens will cause him to jump overboard if he remains unhindered. So he opts to have himself tied to the mast, in order to stop him jumping overboard at a later time.

Let's be a little clearer about Odysseus' thinking. There are two relevant times at which Odysseus must make a choice: he must decide (now) whether to be tied to the mast, and later whether to jump into the sea and drown. If he is tied to the mast now, he will not be able to jump overboard. Also, if he is not tied to the mast, the sirens will drive him mad and make him jump overboard. In a sense, Odysseus' most preferred option is to not be tied to the mast, but then to remain on the boat when in range of the sirens. The problem is that he knows now that his future self will not stay on board if he is unhindered. Odysseus is essentially treating his future self as an agent he does not have full control over, and treating his decision problem as a kind of game: his current self versus his future self. Given what he knows about his future self's preferences (as influenced by the sirens' song), his choice (now) should be that which gets him the best outcome on the assumption that his future self will act in accordance with his future preferences. So he should tie himself to the mast now so that his future self does not have the option to jump into the sea. This is an example of backward induction: a reasoning strategy that is ubiquitous in game theory.

To reiterate: Odysseus' most preferred option would be to not be tied to the mast, but then to refrain from jumping overboard.[18] He does not, however, act on this preference, because he is sophisticated enough to take into account facts about his future choices that are not under his current control. This makes Odysseus a *sophisticated* chooser as opposed to a *naïve* chooser who

[18]Compare: I prefer that both I and my opponent cooperate to both defecting in a Prisoners' Dilemma.

would act on his current preference for not being tied up, and ignore the inevitable future bad consequences of such actions. Sophisticated choice is clearly better than naïve choice: it simply amounts to not ignoring pertinent facts about Your future choices.

Let's go back to the second line of reasoning discussed earlier that seemed to conclude that all the committee members agreed that learning was best. Consider the two extreme committee members Mrs. Black and Mr. White who think that all the black (white) marbles are in urn X. Mrs. Black and Mr. White make up a committee that must collectively make decisions on whether to learn and how to bet. Mrs. Black thinks that learning increases expectation because if X is true betting is the right option and if Y is true not betting is the right option. Mr. White thinks that learning increases expectation because if X is true not betting is the right option and if Y is true betting is the right option. Now consider things from Mrs. Black's point of view. She is reasoning at node 0 about which option she prefers. She knows that if the committee arrives at node 1, she and Mr. White will disagree about what to do, and she is not sure about how such disagreement will be resolved. So at node 0, Mrs. Black cannot discount the possibility that at node 1 the committee will choose not to bet (despite that being a bad option according to Mrs. Black). So what Mrs. Black and Mr. White agree on is not that the group decision should be to choose to learn, but merely that if each of them individually were in control of the future choice, learning would be the right choice. That is, each committee member only believes learning is the right choice because of what the post-learning choices would be if that committee member was in charge. Given each committee member's uncertainty about how future choices will in fact be resolved, it is not clear that the committee will make the 'correct' choice from that member's perspective. And thus, the members do not agree that learning is necessarily the right choice, given what they know about the committee's future choices. And indeed, if the committee members knew that the committee resolved disagreements by using gamma-maximin, they would agree that not learning is the better group decision.

3.2 Modelling assumptions

This section serves two purposes: we introduce some formalism that we'll need later, and we also make clear some assumptions we make in what follows.

First, we shall be using the language of *choice functions* which take as inputs the available options and output the set of choice-worthy or admissible options. For the precise probability case, the standard choice function is

	$C(\mathbf{O_1})$ = Bet; $C(\mathbf{O_2})$ = Bet; $C(\mathbf{O_3})$ = Don't Bet	
	X	Y
Don't Learn	$\langle 3, B; -1, W\rangle$	$\langle 3, B; -1, W\rangle$
Learn	$\langle 3, B; -1, W\rangle$	0

Table 2: Tabular represention of problem in figure 1

the function that takes a set of options and outputs the set of options that maximise expected utility.

In accordance with our sophisticated choice approach, Your current options are the acts available to You at the current time, and not paths through the decision tree or plans of how to act at the current node and all future nodes. Further, from Your current perspective, You take how You anticipate You will act at future nodes to be fixed: to be part of the state space of Your current decision.

Let $\mathbf{O_i}$ be the options available at node i. Let's consider the simple case of the precise example in figure 1. So, for $i = 1, 2, 3$, $\mathbf{O_i}$ is the set "Bet, Don't Bet". And $\mathbf{O_0}$ – the options at node 0 – is "Learn, Don't Learn". The choice function summarises the options that would be chosen at the node, so for node 1, $C(\mathbf{O_1})$ = Bet. And likewise $C(\mathbf{O_2})$ = Bet and $C(\mathbf{O_3})$ = Don't Bet. These facts about future choices should be encoded in the state space of Your decision problem at node 0.[19] As such, this problem should be summarised as per Table 2. The events X and Y have probability 0.5 each. Note that the outcomes in this table are determined by the outcomes that the future choices will lead to. Some of these outcomes are lotteries. The lottery $\langle 3, B; -1, W\rangle$ should be read as follows: 'Win 3 if B, lose 1 if W'. By assumption (recall figure 1), the lotteries in the X column reduce to a sure outcome of 3, since $\Pr(B|X) = 1$, and the lottery in the Y column reduces to a sure outcome of -1, since $\Pr(B|Y) = 0$.[20] Now we again apply the choice rule to this decision problem, and thus determine $C(\mathbf{O_0})$. As noted earlier, learning has a higher expected utility ($EU = 1.5$) than not learning ($EU = 1$).

[19]The implicit assumption here is that You predict Your future self to be rational, i.e. an expected utility maximiser, and to have beliefs and desires that accord with Your current self.

[20]Note in this example Your credences are such that the lottery either has expected value -1 or 3 depending on whether Urn X or Urn Y is being drawn from, since Your conditional credences are extreme. But we present the outcome as the lottery to fit with further examples where You do not have extreme credences.

	$C(\mathbf{O_1}) =$ Bet; $C(\mathbf{O_2}) =$ Don't Bet; $C(\mathbf{O_3}) =$ Don't Bet	
	X	Y
Pay not to learn	$\langle 3-\delta, B; -1-\delta, W\rangle$	$\langle 3-\delta, B; -1-\delta, W\rangle$
Learn	0	0

Table 3: Tabular represention of problem in figure 2 using gamma-maximin

Consider a further example. Table 3 represents Your node-0 decision problem as depicted in figure 2, with the assumption that the gamma-maximin decision rule will be applied at any future choice nodes.

Both of the above examples obscure an important fact about the general case of imprecise sophisticated choice. That is that a choice function may fail to determine a *unique* admissible act. Let's go through the choice problem again, but using a different choice rule. Let's imagine You use the permissive 'non-dominated-set' (NDS) decision rule: all Your choice function does is rule out options that are expectation-dominated. We discuss this rule in more detail in the next section. You are at node 0 reasoning about how Your future selves will behave. Your node 1 self is easy: the acts available – 'Bet', 'Don't bet' – each have precise utility $1-\delta$ and $-\delta$ respectively. So Your node-1 self will choose to bet. Now consider Your node-2 self. Your beliefs at node 2 have dilated, and with them, Your expectations. The expectations for 'Bet' and 'Don't bet' are $[-1, 3]$ and 0, and neither act expectation-dominates the other. So from node 0, You do not know how Your node-2 self will act. The same reasoning goes for node 3. We can thus represent Your node-0 decision problem as per Table 4. Note that choice function C can return a set of acts.

	$C(\mathbf{O_1}) =$ Bet; $C(\mathbf{O_2}) = \{$Bet,Don't Bet$\} = C(\mathbf{O_3})$	
	X	Y
Pay not to Learn	$\langle 3-\delta, B; -1-\delta, W\rangle$	$\langle 3-\delta, B; -1-\delta, W\rangle$
Learn	$\{0, \langle 3, B; -1, W\rangle\}$	$\{0, \langle 3, B; -1, W\rangle\}$

Table 4: Tabular represention of problem in figure 2 for NDS rule

So how should Your node-0 self reason about these unknown future choices? How should learning be valued at node 0? The choice to learn amounts to a fifty-fifty chance of one or another future decision problem, neither of which You know how You will solve. The next section presents our answer

to this question. The presentation may seem complicated, but we maintain that the complications are a feature of imprecise decision theory when it comes to sequential choice. As mentioned, this is one aim of our paper: to make explicit how imprecise decision rules should be characterised in the sequential-choice setting. While others (notably Kadane, Schervish, and Seidenfeld (2008) and Seidenfeld (2004)) have considered how key rules fare in particular sequential-choice problems, with similar findings to our own, no one has provided a clear account of what it means to do backwards induction where numerous incommensurable options are choice-worthy at future nodes.[21]

4 The non-dominated-set rule

The latter part of the discussion in the previous section alluded to an imprecise decision rule that is less 'opinionated' than the gamma-maximin rule. Roughly, if members of Your credal committee disagree in their preference between options, then Your overall preference between these options is simply indeterminate, and there is no saying which will be chosen in a decision between them. This is effectively the non-dominated-set (NDS) rule, otherwise known as *Sen-Walley maximality*. We will define it in more precise terms shortly. The NDS rule is the most permissive decision rule on the table: it would be a serious flaw of any decision rule that it sanctioned performing some action when a dominating action was available (i.e. if the choice rule made permissible something outside of the NDS choice set). This makes NDS interesting for two reasons. First, it is a very plausible decision rule, as it does not contrive a preference between incommensurable options where there is none, so to speak. Second, being the most permissive rule, NDS is pivotal in our discussion. If even the NDS rule made paying to avoid free information obligatory, then this would be a killer blow to imprecise probabilism, since any other less permissive rule would inherit this flaw.

So how the NDS rule fares in 'free evidence' situations is more important that how gamma-maximin fares in those situations. In what follows, we first give a formal definition of the NDS rule (subsection 4.1) and then we prove a result for the rule that relates to the third premise of our trilemma (subsection 4.2). The result is not surprising given the previous work mentioned above, but no-

[21] In fact, some of Seidenfeld (2004)'s less central remarks are at odds with our account of sophisticated choice below: his discussion on pp. 85-6 does not respect the fact that individual 'committee members' (cf. Section 3.1) do not have control over the options that will be chosen at future nodes, that this is rather within the purview of the whole committee.

one has yet given a treatment as general as ours, nor as explicit with respect to sequential-choice reasoning.[22] We provide such a general treatment and articulate the correct albeit controversial approach to sophisticated choice.

4.1 Formal definition

In formal terms, the NDS rule can be defined in terms of its 'choice set' (i.e. the set of admissible options that it returns) as follows:

$$N(\mathbf{O}) = \{A_j \in \mathbf{O} : \forall A_k \in \mathbf{O}, \neg[A_j <_{EU} A_k]\}$$

This needs some unpacking. The meaning of $A_j <_{EU} A_k$ that accords best with the current literature is as follows: option A_j has lower expected utility than option A_k according to *all* probability distributions in Your representor. In such a case, A_j is strongly dominated by A_k and is thus not choice-worthy. This definition of $<_{EU}$ will not quite do, however, given the complications that can arise for the imprecise probabilist making sequential decisions, that we canvassed at the close of the previous section.

In order to spell out an adequate definition of $<_{EU}$ and thus the NDS rule for sequential contexts, it helps to return to our example of the imprecise probabilist facing the decision problem in figure 2. In accordance with backwards induction, we determine what will be chosen at the later nodes: it is clear enough that $C(\mathbf{O_2}) = C(\mathbf{O_3}) = \{\text{Bet, Don't Bet}\}$ by the NDS rule, since neither betting nor not betting *EU*-dominates the other.[23] Indeed, thinking in terms of our credal committee: Mrs. Black and Mr. White disagree on what choices at nodes 2 and 3 maximise expected utility. As before, $C(\mathbf{O_1}) = \text{Bet}$, since all committee members agree that this option at node 1 has higher expected utility. The upshot is that the problem facing the NDS agent (the credal committee) at the initial node is in fact the problem represented in

[22] In particular, see Kadane, Schervish, and Seidenfeld (2008) Sequential Decision Problem 2 for a decision problem similar to our example in figure 2. Seidenfeld (2004) gives a general treatment of the free-evidence problem for both gamma-maximin and Levi's *e-admissibility* imprecise decision rule. The latter rule is very similar to NDS: An option is *e-admissible* (and therefore in the choice set, unless a secondary criterion is invoked) just in case it has maximal expected utility for at least one probability-utility pair in the agent's belief and desire representor(s). The set of e-admissible options is always a subset of the NDS choice set, and in ordinary cases, where sets of probabilities and utilities are closed and convex, the set of e-admissible options is in fact identical to the NDS choice set; see Schervish et al. (2003). Our analysis below differs from Seidenfeld's in that we focus rather on the NDS rule itself, and we show how this rule needs to be extended to accommodate possibilities that may arise in the sequential-choice setting.

[23] Recall that $C(\mathbf{O_i})$ corresponds to the admissible choices at node i in our example problem.

Table 4. Everyone in the committee sees, at node 0, that at nodes 2 and 3 there will be disagreement amongst the committee members, and either betting or not betting could be chosen at those locations. So the choice set for node 2 contains both the bet and don't bet options, likewise for node 3. Thus, the outcome of opting to learn – which is a mixture of the outcomes of nodes 2 and 3 – amounts to a mixture of two sets of two future acts.

So our NDS agent faces the decision problem depicted in Table 4: the complicated case involving outcomes with sets of acts and thus sets of utilities. For each probability distribution $\Pr_i \in \mathcal{P}$ (or each committee member i), the evaluation of an act A_j (here learning) may thus amount to a *set* of expected utilities.

Let us now continue with unpacking the statement $A_j <_{EU} A_k$:

$$\begin{aligned} A_j <_{EU} A_k \text{ IF} \\ \forall \Pr_i \in \mathcal{P} \ [\mathbf{EU}_i(A_j) <_P \mathbf{EU}_i(A_k)] \end{aligned}$$

$\mathbf{EU}_i(A_j)$ is the expected utility of act A_j according to probability function $\Pr_i$. It can be a set of values if act A_j is not the terminal node of the decision tree, because of Your uncertainty about Your post-A_j choices. Of course, the problem now shifts to defining $<_P$. All that has been stipulated about $<_P$ is that, in the case where both $\mathbf{EU}_i(A_j)$ and $\mathbf{EU}_i(A_k)$ are singletons, $<_P$ amounts to the simple 'is less than' relation for numbers. Thus, when A_j and A_k lead to outcomes with precise expected utilities for each $\Pr_i$, $A_j <_{EU} A_k$ has its standard meaning. Given that the standard NDS rule appeals to a *strictly* greater than relation, it makes sense to define $<_P$ as follows:

$$\begin{aligned} \mathbf{EU}_i(A_j) <_P \mathbf{EU}_i(A_k) \text{ IF} \\ \sup \mathbf{EU}_i(A_j) < \inf \mathbf{EU}_i(A_k) \end{aligned}$$

The relation $<_P$ is sometimes known as the relation of *interval dominance*. This makes sense because it holds exactly when one utility interval is entirely above another. Note that it does behave appropriately in the precise limit.[24]

[24]cf. Levi's e-admissibility rule, which we denote $L(\mathbf{O})$. Recall from footnote 22 that options are in $L(\mathbf{O})$ if and only if the option is maximal for at least one probability distribution in Your representor. To account for the complications that arise in the sequential-choice setting (outcomes that are sets of possible choices), we could augment Levi's rule as follows: for an option to be e-admissible, there must be at least one probability distribution in Your representor such that no other option interval dominates that option according to that probability:

$$L(\mathbf{O}) = \{A_i \in \mathbf{O} : \exists \Pr \in \mathcal{P} \text{ s.t. } \forall A_k \in \mathbf{O} \ \neg[\mathbf{EU}_{\Pr}(A_k) \geq_P \mathbf{EU}_{\Pr}(A_i)]\}$$

We can now finish working through the decision problem in figure 2 using our fully-worked-out NDS rule. We have already established, via backwards induction, that the decision problem at the initial node is given in Table 4. The question is whether learning *EU*-dominates not learning, or vice versa. Consider just one of the extreme probability distributions – that of committee member Mrs. Black – $\Pr_1$, where $\Pr_1(B|X) = 1$ and $\Pr_1(B|Y) = 0$. *For this probability distribution alone*, the lotteries in Table 4 will thus be evaluated as per Table 5 below. Note that the lottery outcome in the X column is evaluated as 3 (or $3 - \delta$), since $\Pr_1(B|X) = 1$. Thus the lottery outcome in the Y column is evaluated as -1 (or $-1 - \delta$).

	$C(\mathbf{O_1}) =$ Bet; $C(\mathbf{O_2}) = \{$Bet,Don't Bet$\} = C(\mathbf{O_3})$	
	X	Y
Pay not to Learn	$3 - \delta$	$-1 - \delta$
Learn	$\{0, 3\}$	$\{0, -1\}$

Table 5: Tabular represention of problem in figure 2 for the NDS rule, for $\Pr_1$ (Mrs. Black)

To work out Mrs. Black's evaluation of the acts in Table 5 (i.e. learning versus paying not to learn), we need to generalise the idea of expected value to sets of utilities. An informal characterisation of expected value is 'the sum of the probability weighted utilities'. So it's the probability of the state times the utility of the act in that state, summed over the states. To evaluate the act 'Learn' in Table 5 above, we need a 'sum of probability weighted utility sets'. This means characterising what it means to weight a set of utilities by a probability, and then what it means to add sets of probability-weighted utility together. In other words, we need to evaluate:

$$\mathbf{EU}_1(\text{Learn}) = 0.5 \times \{3, 0\} + 0.5 \times \{-1, 0\}$$

We will need to give an appropriate gloss on what $\times$ and $+$ mean in this context. We suggest that $p \times A = \{pa : a \in A\}$. That is, the multiplication is done to each element of A. As for $A + B$, we take this to mean $\{a + b : a \in A, b \in B\}$. This is the full set of possible sums of elements of A and B.[25] This proposal for how to treat sums and products of sets of values can

This differs from NDS only in the order of the quantifiers. Clearly e-admissible options are not *EU*-dominated; rather $L(\mathbf{O}) \subseteq N(\mathbf{O})$.

[25]If A, B are intervals (i.e. convex), and we interpret p as the degenerate interval $[p, p]$, then this proposal gives the same result as *interval arithmetic* (Moore, Kearfott, and Cloud

be motivated in the same way we motivated the NDS rule: this is the most permissive plausible proposal for how to deal with arithmetic operations on sets of values. So this approach is pivotal in the same sense that the NDS rule is. Returning to the example, this gives:

$$\mathbf{EU}_1(\text{Learn}) = \{-0.5, 0, 1, 1.5\}$$

Paying not to learn has precise utility $\{1 - \delta\}$. Thus it is *not* the case that $\mathbf{EU}_1(\text{Learn}) <_P \mathbf{EU}_1(\text{Pay not to learn})$, nor vice versa.[26] Accordingly, just by looking at Mrs. Black's evaluation of the options (let alone the evaluations of her fellow committee members) we see that neither learning nor paying not to learn *EU*-dominates the other. So at the initial node, both options are admissible by the NDS rule, i.e. $C(\{\text{Learn}, \text{Pay not to learn}\}) = \{\text{Learn}, \text{Pay not to learn}\}$. The NDS rule therefore sanctions paying not to learn, but in this case, at least, the rule does not *require* or *mandate* paying for ignorance.[27] This is an improvement on the gamma-maximin rule.

4.2 A positive result

While we have seen that paying not to learn may be an admissible option, according to the NDS rule, there is some good news: for this pivotal imprecise decision rule, learning free evidence is always itself an admissible option. That is, it is never the case that not learning or else paying not to learn is uniquely admissible or obligatory, when learning free evidence is also an available option.

The structure of the proof follows our working of the decision problem above, but the idea here is to show that this analysis generalises to all decision scenarios of the following form: Your beliefs may be imprecise, i.e. Your belief representor, $\mathcal{P}$, may consist of more than one probability distribution. (Your utility function over basic outcomes is nonetheless precise.) You face a decision as to whether to pursue free evidence, *sensu* Good (1967). As before, it is a question of whether to make a choice between a set of acts now, or else make the choice between the same set of acts (with basic outcomes unchanged) after learning some evidence. We assume, for simplicity, that the decision to be made now or after learning is not itself a sequential decision

2009). Since the $>_P$ relation depends only on the biggest and smallest members of A, B, it treats nonconvex sets in the same way as their smallest convex cover.

[26] For $\delta < 1.5$.

[27] Levi's e-admissibility rule in fact gives the same result here: there is at least one probability function in the representor, the one where $Pr(B|X) = 0.5$, for which both learning and not learning have maximal expected utility in the sense stated in footnote 24, so both options are in the choice set.

problem. For rhetorical reasons, the option of not learning is adjusted to include a payment of some small δ. So You can either learn and then make a decision, or You can pay to not learn and make the decision from a state of ignorance.

This is the claim to be proven, with reference to the general decision problem just described:

> For the imprecise agent abiding by the NDS rule, learning free evidence is *always* an admissible option. That is, not learning or else paying not to learn free evidence is never uniquely admissible, when learning is also an available option.

We start by summarising Good's result. Let's say we have acts whose outcomes depend on which of the H_ms is true. Call the acts the A_js. We are trying to work out whether it pays in expectation to learn which E_k is true. The H_ms partition the state space, as do the E_ks. The expectation for not learning is:

$$\max_j \sum_m \Pr(H_m) U(A_j(H_m))$$

That is, You know You will pick the act that maximises Your utility after having chosen not to learn. Using the fact that $\Pr(H_m) = \sum_k \Pr(E_k) \Pr(H_m|E_k)$ and rearranging the order of summation, we have that this is equal to:

$$\max_j \sum_k \sum_m \Pr(E_k) \Pr(H_m|E_k) U(A_j(H_m)) \tag{2}$$

Now consider learning some E_k. Expectation after learning E_k is:

$$\max_j \sum_m \Pr(H_m|E_k) U(A_j(H_m))$$

So expectation for learning is the probability weighted sum of these expectations:

$$\sum_k \Pr(E_k) \max_j \sum_m \Pr(H_m|E_k) U(A_j(H_m))$$

We can move the $\Pr(E_k)$ inside the summation so this is equal to:

$$\sum_k \max_j \sum_m \Pr(E_k) \Pr(H_m|E_k) U(A_j(H_m)) \tag{3}$$

These expressions for learning (3) and not learning (2) differ only in the order of the summation and maximisation. So all we need to do to prove the

theorem is show that, for any function $f(j,k)$, we have

$$\sum_k \max_j f(j,k) \geq \max_j \sum_k f(j,k)$$

Then if we consider $f(j,k) = \sum_m \Pr(E_k) \Pr(H_m|E_k) U(|A_j(H_m))$ we have the result we wanted. So let's consider the j that maximises $\sum_k f(j,k)$, call it j_0. Clearly $\max_j f(j,k) \geq f(j_0,k)$ for any k. Thus

$$\sum_k \max_j f(j,k) \geq \sum_k f(j_0,k) = \max_j \sum_k f(j,k)$$

That, in brief, is Good's theorem.[28]

Returning to our own decision scenario: Let **O** denote the set of acts at the initial node: 'Learn' and 'Don't learn'/'Pay not to learn'. Let **A** denote the set of acts for the decision problem that needs to be solved: 'Bet'/'Don't bet' (i.e. the acts that are the rows in Table 1). The possible evidence forms a partition **E**: 'Urn X' or 'Urn Y'.

Now consider the plight of the imprecise agent employing our refined version of the NDS rule, as outlined in Section 4.1. Assume Your representor, $\mathcal{P}$, consists of probability functions (a. k. a. committee members) $\Pr_1, \Pr_2, \Pr_3, \ldots$ Take any one of these probability functions $\Pr_i$. Quite simply, for this committee member i, we know from Good's theorem itself that, if this committee member were a one-person committee, learning is at least as good as not learning. That is not necessarily the case, given our assumption of imprecise probabilism – there may be other members in the committee. Nonetheless, the acts in **A** that committee member i deems maximal at each of the respective future choice nodes associated with learning, corresponding to each E_k that may be learnt, will feature in the choice sets for these nodes (since, by assumption, the NDS rule is applied at these nodes too.) Thus, when we evaluate learning versus not learning for any committee member i, the probability weighted average of *utility sets* for learning, as discussed above, will be a set of utilities that includes a utility that is guaranteed, by Good's theorem, to be at least as good as the utility of not learning. So for *any* committee member i, not learning does not (strictly interval-)dominate learning. Therefore, not learning (and thus paying not to learn) does not EU-dominate learning, and so learning is always admissible, for our fully worked out version of the NDS rule. The basic idea is that if an act is maximal for some agent, it cannot be dominated.

[28]Good's result also holds in the infinite case. Our result also seems to go through for the infinite case, but we hesitate to assert that since we have not yet looked carefully at whether 'integrals of sets of values' are well behaved.

For clarity, we will restate this proof sketch now more formally. For any probability function $\Pr_i$, each possible evidence $E_k \in \mathbf{E}$ is associated with an act with maximum precise expected utility amongst the set of available acts $\mathbf{A}$. Call this act a_{ik}. (Assume, with no loss of generality, that there is a unique such act with maximum precise expected utility for each $\Pr_i$ and each E_k.)

There is also an act in $\mathbf{A}$ that has maximum expected utility, in the unconditional sense (before learning), according to $\Pr_i$. Again, assume, with no loss of generality, that there is a unique such act, which we label A_{max_i}.

We know, by Good's theorem, that:

$$\sum_k \Pr_i(E_k) \times U_i(a_{ik}(E_k)) \geq \sum_k \Pr_i(E_k) \times U_i(A_{max_i}(E_k))$$

Here U_i is the utility function that is paired with probability function $\Pr_i$.

The right-hand term above amounts to the value of not learning for this probability function $\Pr_i$ in the representor, i.e. $\mathbf{EU}_i$(Don't learn).

The left-most term does not amount to the value of learning with respect to $\Pr_i$, however, because we have not taken into account sophisticated choice. The value of learning for each $\Pr_i$ depends on what would be the sets of admissible acts if each respective E_k was learnt (the $C_k(\mathbf{A})$). For each E_k, these choice sets will at least include the union of all the acts that have maximum expected utility, conditional on E_k, for the probability distributions $\Pr_i$ in $\mathcal{P}$:

$$\begin{aligned} C_k(\mathbf{A}) &\subseteq a_{1k} \cup a_{2k} \cup \ldots \\ &\subseteq \bigcup_i a_{ik} \end{aligned}$$

The set of expected utilities that $\Pr_i$ assigns to this choice set under the assumption that E_k obtains is then

$$U_i(a_{1k}(E_k)) \cup U_i(a_{2k}(E_k)) \cup \ldots = \bigcup_i U_i(a_{ik}(E_k))$$

So the value of learning according to the probability function $\Pr_i$:

$$\mathbf{EU}_i(\text{Learn}) = \sum_k \Pr_i(E_k) \times \bigcup_i U_i(a_{ik})$$

As noted in Section 3, we take the above expression to equate to the set of all combinations of probability weighted sums of utilities, given the utility sets associated with each $C_k(\mathbf{A})$. Therefore, *one* of the utility values in the set $\mathbf{EU}_i(\text{Learn})$ is $\sum_k \Pr_i(E_k) \times U_i(a_{ik})$. That is, the expected utility of choosing according to the expectations of $\Pr_i$ for each k.

Recall that, by Good's theorem, this expression is greater than or equal to the value of not learning, for the probability function in question. Thus, given our definition of $<_P$, it is NOT the case that

$$\mathbf{EU}_i(\text{Learn}) <_P \mathbf{EU}_i(\text{Don't learn})$$

Therefore, it is not the case that

$$\text{Learn} <_{EU} \text{Don't learn}$$

That is, learning is always admissible (given a choice of learning and not learning) by our refined NDS rule.[29]

4.3 Return to the trilemma

If the only plausible decision rule was the NDS rule, the trilemma stated in Section 2 would stand. We have seen that this rule sanctions paying to avoid free evidence, so $P3$ is true. $P1$ and $P2$ are also apparently true, but the three together yield a contradiction. The positive result proved above, however, suggests a possible resolution: This is to deny $P1$ in favour of a weaker premise, $P1'$, stating:

$P1'$ Paying to avoid free evidence, or even avoiding free evidence, should never be uniquely admissible, when pursuing free evidence is also an available option.

Note that in the precise case, we can satisfy the stronger principle that learning free evidence before making a decision has to be *at least as good* as not learning. This means the precise agent never pays to avoid free evidence. The above principle amounts to the weaker claim that learning *cannot be worse*. Since the precise Bayesian has complete preferences, these two principles amount to the same thing in the precise case. In the imprecise case, however, they are importantly different. We have seen that the refined NDS rule satisfies $P1'$, but the gamma-maximin rule does not.

[29] One can also infer from the above that for Levi's rule, too, learning is always admissible. Indeed, not just for one but for *all* probability functions $\Pr_i$, the supremum of the expected utility values for learning is greater than the expected utility for not learning.

5 Broader considerations of dynamic coherence

Could we do better than the NDS rule? Should we be looking for a decision rule that never even permits paying to avoid free evidence? We think the answer is no. There's a sense in which decision rules like NDS[30] cannot be improved on.

If one wants to advocate some form of non-probabilistic epistemology, and if one wants to link this up with decision theory, then some axiom of the standard expected utility representation theorems must be denied. The plausible candidates for denial are the ordering postulate and the independence postulate. The NDS rule violates ordering, while the gamma-maximin rule upholds ordering but violates independence.

Any denial of independence leads to a pragmatic kind of dynamic incoherence: effectively refusing free money. Indeed, the dynamic incoherence can be seen as a sort of 'information aversion' or paying to avoid free evidence (Seidenfeld 1988; Wakker 1988; Al-Najjar and Weinstein 2009). This is easily shown via the decision problem in figure 3. Let A, B, C stand for lotteries and ApC stand for the mixed lottery 'A if Z, C otherwise' where Z has probability p. Let $\succ$ be Your strict preference – that is, $\succ$ is asymmetric and transitive. Independence is the principle that states that $A \succ B$ if and only if $ApC \succ BpC$ for all C. A violation of independence is thus a set of lotteries A, B, C where $A \succ B$ but it is not the case that $ApC \succ BpC$. Assume a *strong* violation of independence, i.e. one involving a preference reversal: $BpC \succ ApC$. A sophisticated agent with the above independence-violating preferences reasons about the decision in figure 3 as follows (Seidenfeld 1988; Steele 2010): You start by looking at the 'terminal' decisions: those decisions with no further choices following them. In this case, those are the choices at nodes 1 and 2. In each case, the arrow to the right highlights the hypothetically chosen option. Both these choices can be read off the preferences: $A \succ B$, so A is chosen at node 1, but $BpC \succ ApC$, so BpC is chosen at node 2. Now consider the choice from node zero. Since You know that your node 1 choice will be to choose A, the 'up' option of node 0 is effectively the gamble ApC. The 'down' option of node 0 is a choice between the two gambles ApC and BpC. Since You prefer BpC to ApC, you prefer to go down at node 0. Since the preference is strict, there is some small amount of money, δ, that You would pay to go down if offered this decision problem. Recall that p is the probability of event Z. Note that choosing up is effectively choosing to

[30] We noted in the previous section that our results concerning the NDS rule also apply to Levi's e-admissibility rule (cf. Seidenfeld (2004).)

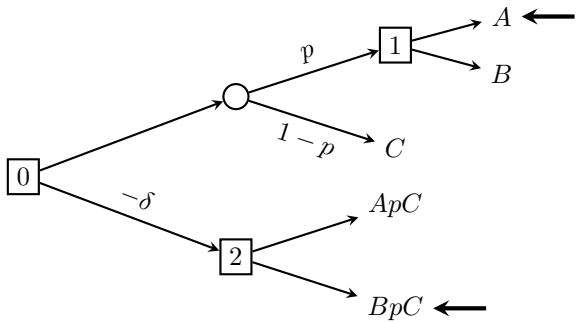


Figure 3: A decision tree that yields information aversion (after Seidenfeld (1988))

learn whether Z is true before making Your choice. Since You prefer down to up, You would pay to avoid learning whether X has occurred. The only thing needed to get this example going was a strong violation of independence.[31]

We advocate, instead, dropping the ordering postulate: the induced preference can be incomplete. As we have seen, such decision rules also sometimes suffer from issues with information aversion. However, they merely permit – never mandate – paying to avoid free evidence. This is an advantage over independence-violating decision rules. See also Bradley and Steele (2014a) for further discussion of permissision and obligation in the context of sure loss and sequential choice.

6 Concluding Remarks

Let us return to the trilemma that we began with:

$P1$ *Paying to avoid free evidence is irrational.*

$P2$ *Incomplete preferences are not irrational, and can be represented by sets of probabilities and utilities. In other words, imprecise probabilism is permissible.*

$P3$ *All plausible decision theories for handling imprecision sanction paying to avoid free evidence.*

Of course, one could always deny $P2$, and claim that imprecise probabilism is evidently irrational. But we think there are compelling reasons for beliefs

[31]Epstein and Le Breton (1993)'s theorem also dashes hopes for theories that violate Independence but retain Ordering. The theorem effectively shows that the imprecise probabilist cannot keep Ordering and also avoid information aversion.

being imprecise, and so we consider this 'way out' to be a last resort (but see Al-Najjar and Weinstein (2009) for a different opinion).

Alternatively, one could deny that 'free evidence' has any natural meaning. This could be reasonable: 'free' only makes sense with respect to some notion of value, and notions of value seem intimately tied to theories of decision. One might say that evidence is free by the lights of some decision theory just in case You would not pay to avoid this evidence if You were abiding by the theory. In terms of our trilemma, this move denies $P3$—that decision theories for handling imprecision allow paying to avoid free evidence—by claiming that the evidence is not free, by definition. Of course, the spin-off of this move is that $P1$ is rendered vacuous. A resolution of this sort might be appealing to those who already regard free evidence a merely technical term in the precise context.

We do not consider this a reasonable response, since Good's definition of free evidence seems intuitive, and applicable to a wide class of decision theories. Recall that evidence is free by the standards of Good's proof if learning this evidence does not otherwise change Your decision problem. That is, You update Your beliefs according to conditionalisation, but the available options $A_1, \ldots, A_s$, and the utilities of the outcomes associated with these options remain the same, whatever evidence is learnt.[32]

One might object to $P1$ on the grounds that there are cases of intuitively free evidence that You might reasonably want to avoid. For instance, imagine You are trying to decide between reading *Murder on the Orient Express* and going for walk, and then someone offers to tell You who killed Mr. Ratchett before You make the decision. Common sense tells us that it may be disadvantageous for You to accept this evidence, even if it comes with no explicit charge, which seems strangely at odds with Good's theorem.

Closer inspection of the above case reveals that the evidence, while not incurring a monetary or other material cost, is not free in the Good sense because it may change the outcomes of the reading option: reading the novel when You know who did it is not the same act as reading the novel when You are one step behind Poirot and his little grey cells. That is to say that there is intrinsic value to Your current state of ignorance with respect to the identity of the murderer: you enjoy your reading of the novel more while not knowing who did it. Kadane, Schervish, and Seidenfeld (2008) discuss a similar case

[32]Of course, in many cases it is beneficial for the precise probabilist to pursue costly evidence; it depends on the details of the case at hand. Good's proof is interesting, however, because it establishes the general claim: that it is *never* a bad thing to pursue free evidence.

(Example 12). States of belief having intrinsic value offer well-known exceptions to various Bayesian theorems, for example Dutch book arguments for additivity or conditionalisation.

What we have done in this paper is examine premises $P1$ and $P3$, with their meanings taken at face value. We have shown that something stronger than $P3$ in fact holds – it is not just the 'extant' decision theories that permit paying to avoid evidence, but in fact any plausible decision theory for handling imprecision permits paying to avoid free evidence in some cases. On the positive side, however, we have refined the consequence of 'paying to avoid evidence'. Certain decision theories – those akin to the NDS rule[33] – do indeed permit paying to avoid free evidence in 'problem cases' that apparently all involve dilation, but these theories never have avoiding free evidence as the uniquely admissible option when learning is available.

So, if we are willing to revise premise $P1$ in favour of $P1'$, then there is no longer a contradiction. Note that $P3$ can also be revised to $P3'$ so that it is more informative:

$P1'$ *Paying to avoid free evidence, or even avoiding free evidence, should never be uniquely admissible, when pursuing free evidence is also an available option.*

$P2$ *Incomplete preferences are not irrational, and can be represented by sets of probabilities and utilities. In other words, imprecise probabilism is permissible.*

$P3'$ *For a certain class of decision rules akin to the Sen-Walley Maximality rule, learning free evidence is always admissible, but paying to avoid free evidence may sometimes be admissible too. There are no plausible decision rules that fare better than this, in the sense of never permitting paying to avoid free evidence.*

This is, in our opinion, the best weakening or 'way out' of the original trilemma. It gives us the conclusion that the NDS rule is as good as it gets for the imprecise probabilist when it comes to negotiating free evidence.

[33]The NDS rule is the most 'permissive' decision rule, but we have noted that rules at least as permissive as Levi's e-admissibility rule behave similarly to NDS with respect to free evidence.

References

Augustin, Thomas et al., eds. (2014). *Introduction to Imprecise Probabilities*. John Wiley and Sons.

Bradley, Seamus and Katie Steele (2014a). "Should Subjective Probabilities be Sharp?" *Episteme* 11, pp. 277–289.

——— (2014b). "Uncertainty, Learning and the "Problem" of Dilation". *Erkenntnis* 79, pp. 1287–1303.

Epstein, Larry G. and Michel Le Breton (1993). "Dynamically Consistent Beliefs Must Be Bayesian". *Journal of Economic Theory* 61, pp. 1–22.

Gärdenfors, Peter and Nils-Eric Sahlin (1982). "Unreliable probabilities, risk taking and decision making". *Synthese* 53, pp. 361–386.

Good, Irving John (1967). "On the principle of total evidence". *British Journal for the Philosophy of Science* 17, pp. 319–321.

Grünwald, Peter D. and Joseph Y. Halpern (2004). "When ignorance is bliss". In: *Proceedings of the Twentieth Conference on Uncertainty in AI*, pp. 226–234.

Halpern, Joseph Y. (2003). *Reasoning about Uncertainty*. MIT press.

Huttegger, Simon (2013). "Learning Experiences and the Value of Knowledge". *Philosophical Studies* 171, pp. 279–288.

Joyce, James M. (2005). "How Probabilities Reflect Evidence". *Philosophical Perspectives* 19, pp. 153–178.

Kadane, Joseph B., Mark J. Schervish, and Teddy Seidenfeld (2008). "Is ignorance bliss?" *Journal of Philosophy* CV, pp. 5–36.

Kaplan, Mark (1996). *Decision Theory as Philosophy*. Cambridge University Press.

Levi, Isaac (1974). "On Indeterminate probabilities". *Journal of Philosophy* 71, pp. 391–418.

——— (1980). *The Enterprise of Knowledge*. The MIT Press.

——— (1986). *Hard Choices: Decision Making Under Unresolved Conflict*. Cambridge University Press.

——— (1991). "Consequentialism and Sequential Choice". In: *Foundations of Decision Theory*. Ed. by Michael Bacharach and Susan Hurley. Blackwell, pp. 92–122.

——— (1999). "Value commitments, value conflict and the separability of belief and value". *Philosophy of Science* 66, pp. 509–533.

Maher, Patrick (1992). "Diachronic Rationality". *Philosophy of Science* 59, pp. 120–141.

Miller, David (1994). *Critical Rationalism: A Restatement and Defence*. Open Court.

Moore, Ramon E., R. Baker Kearfott, and Michael J. Cloud (2009). *Introduction to Interval Analysis*. Society for Industrial and Applied Mathematics.

Al-Najjar, Nabil I. and Jonathan Weinstein (2009). "The Ambiguity Aversion Literature: A Critical Assessment". *Economics and Philosophy* 25, pp. 249–284.

Pedersen, Arthur Paul and Gregory Wheeler (2014). "Demystifying Dilation". *Erkenntnis* 79, pp. 1305–1342.

Raiffa, Howard and Robert Schlaifer (1961). *Applied Statistical Decision Theory*. Harvard Business School.

Ramsey, Frank P. (1990). "Weight of the Value of Knowledge". *British Journal for the Philosophy of Science* 41, pp. 1–4.

Savage, L.J. (1972 [1954]). *The Foundations of Statistics*. 2nd ed. Dover.

Schervish, Mark J. et al. (2003). "Extensions of Expected Utility Theory and Some Limitations of Pairwise Comparisons". In: *Proceedings of ISIPTA 2003*, pp. 496–510.

Seidenfeld, Teddy (1988). "Decision theory without "independence" or without "ordering". What's the difference?" *Economics and Philosophy*, pp. 267–290.

——— (2004). "A contrast between two decision rules for use with (convex) sets of probabilities: Gamma-maximin versus E-admissibility". *Synthese* 140, pp. 69–88.

Seidenfeld, Teddy, Joseph B. Kadane, and Mark J. Schervish (1989). "On the shared preferences of two Bayesian decision makers". *The Journal of Philosophy* 86, pp. 225–244.

Seidenfeld, Teddy, Mark J. Schervish, and Joseph B. Kadane (1995). "A Representation of Partially Ordered Preferences". *Annals of Statistics* 23, pp. 2168–2217.

Seidenfeld, Teddy and Larry Wasserman (1993). "Dilation for sets of probabilities". *Annals of Statistics* 21, pp. 1139–1154.

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

Steele, Katie (2010). "What are the minimal requirements of rational choice? Arguments from the sequential setting". *Theory and Decision* 68.4, pp. 463–487.

Sturgeon, Scott (2008). "Reason and the grain of belief". *Noûs* 42, pp. 139–165.

Wakker, Peter (1988). "Nonexpected Utility as Aversion to Information". *Journal of Behavioural Decision Making* 1, pp. 169–175.

Walley, Peter (1991). *Statistical Reasoning with Imprecise Probabilities*. Vol. 42. Monographs on Statistics and Applied Probability. Chapman and Hall.