Theory and Decision (2023) 94:1–33
https://doi.org/10.1007/s11238-022-09883-y

Desirability relations in Savage's model of decision making

Dov Samet[1] · David Schmeidler[1]

Accepted: 20 February 2022 / Published online: 7 March 2022

Abstract

We propose a model of an agent's probability and utility that is a compromise between Savage (The foundations of statistics, Wiley, 1954) and Jeffrey (The Logic of Decision, McGraw Hill, 1965). In Savage's model the probability–utility pair is associated with preferences over acts which are assignments of consequences to states. The probability is defined on the state space, and the utility function on consequences. Jeffrey's model has no consequences, and both probability and utility are defined on the same set of propositions. The probability–utility pair is associated with a desirability relation on propositions. Like Savage we assume a set of consequences and a state space. However, we assume that states are *comprehensive*, that is, each state describes a consequence, as in Aumann (Econometrica 55:1–18, 1987). Like Jeffrey, we assume that the agent has a preference relation, which we call *desirability*, over events, which by definition involves uncertainty about consequences. For a given probability and utility of consequences, the desirability relation is presented by conditional expected utility, given an event. We axiomatically characterize desirability relations that are represented by a probability–utility pair . We characterize the family of all the probability–utility pairs that represent a given desirability relation.

Keywords Savage's model · Jeffrey's model · Desirability · Preferences · Utility · Consequences · Subjective probability

1 Introduction

1.1 Savage vs. Jeffrey

Savage (1954) and Jeffrey (1965) each studied a model in which a preference relation is associated with subjective probability. We begin by describing the two

✉ Dov Samet
dovsamet@gmail.com

David Schmeidler
davidschmeidler@gmail.com

1 Coller School of Management, Tel Aviv University, Tel Aviv, Israel

models and delineate the model that we study in this paper which is a compromise between Savage's and Jeffrey's models.

The building blocks of Savage's theory are consequences, acts, and states of the world. The agent in this theory faces various acts each of which may result in some possible consequences. The agent is uncertain about which consequence will be realized by her acts. This uncertainty is modeled by specifying a set of states of the world. Each state is a rich enough description of the world that resolves these uncertainties, namely, each state determines the consequence of each one of the acts. Thus, the uncertainty about the consequence of an act is translated into uncertainty about the state of the world. Probability in this theory is defined on events, that is, subset of states, while utility is defined over consequences. The expected utility of acts defines a preference relation over acts. These preferences satisfy certain axioms, and fulfilment of these axioms makes it possible to find a probability–utility pair that gives rise to the preferences.

The primitive notion of Jeffrey's model is a set of propositions. Two measures are defined on this set: A signed measure and a probability measure. The signed measure of a proposition is thought of as the total utility it provides, although the theory does not explicitly define utility. The ratio of the signed measure and the probability measure is called the *desirability function*. The desirability of propositions naturally defines a preference relation over propositions. We call such a relation a *desirability relation*. The desirability relation satisfies certain axioms, and when a preference relation on propositions satisfies these axioms then the relation can be shown to be defined by the ratio of the two measures as described above.

Here we propose a model which integrates elements from the two models that at first glance seem to have nothing in common. First we show what is indeed common to them. The set of proportions is a Boolean algebra, and so is the set of subsets of a state space. Thus, disregarding the mathematical details, we can think of the events in Savage's model as being Savage's propositions.[1] Hence, the signed measure and the probability measure in Jeffrey's model can be thought of as measures defined on events. We said before that the signed measure in Jeffrey's model can be though of as total utility delivered by the proposition. Now that we think of propositions as being events, we can define a utility function, defined on the state space, as the derivative of the signed measure with respect to the probability. This way, the signed measure of an event becomes the integral of the utility over the event, which is indeed the total utility provided by the event. The ratio of the signed measure and the probability, is the expected utility given the event. At this point the affinity between the two models breaks down. The utility thus derived is defined on states, while in Savage's model it is defined on consequences. In Jeffrey's model there are no consequences and therefore no acts.

In our model we keep the state space with the Boolean algebra of its events and the probability on events as in Savage's model. As we have seen this also agrees with Jeffrey's model though we gain the concept of states of the world that is

[1] This is done with almost no loss of generality as by the Stone Theorem, Stone (1936), every Boolean algebra is isomorphic to the Boolean algebra of a family of subsets.

exclusive to Savage's model. But we also want to keep consequences and utility defined on them as in Savage's model, and at the same time have a relation of desirability on events as in Jeffrey's model.

The difficulty in combining together all these elements is the nature of states of the world in Savage's model. A state is a description of the world that enables us to determine the consequence of each act. Thus, the description of the world in a state does not and should not include a specific consequence and it is consistent with any consequence. But, then, we cannot talk about the expected utility given an event, when the utility is defined on consequences and the states in the event are not related to any specific consequences. Therefore we cannot talk also about the desirability of an event or about a desirability relation.

If, however, we fix one act in Savage's model, then each state, combined with the act, determines a consequence. We can now think of the state as being *comprehensive*, that is, it is a full description of the world, *including a consequence*. With this it is possible to talk about the expected utility given an event, and hence to define a desirability function and a desirability relation on events in the set of comprehensive states, à la Jeffrey. We formulate several axioms on a preference relation on events that guarantee that it is a desirability relation defined by a probability on events and utility over consequences.

1.2 Comprehensive states

Comprehensive states, that is, states that also describe a consequence, should come as natural objects when we consider an interaction of decision makers, rather than an a acto of a single decision maker. There is nothing in Savage's model that excludes the possibility of other agents being there whose acts may influence the consequence of the agent's act. In other words, the theory can be applied even when the agent is a player in a game. In this case, however, in order to determine the consequence of the player's chosen act, the description of the world should specify the acts of the other players.

An extension of a very well known example, taken from Savage (1954), illustrates this. Our agent considers the problem of breaking an egg and adding it to a bowl with five eggs previously broken for making an omelette. There are two states of the world: the sixth egg can be good or rotten. The omelette maker has three possible acts: breaking the egg into the bowl, breaking it into a saucer for inspection, or throwing it away. The consequences describe the number of eggs in the omelette and the need to wash the saucer if it was used.

Imagine now another agnet, the egg seller, who sold the sixth egg to the omelette maker The egg seller has new good eggs and old rotten eggs, and he can discern between them. The egg seller is facing two acts: selling a good egg or a rotten one. The states of the world in the omelette maker's model that specify whether the egg is good or rotten, describe the acts available to the egg seller.

Consider now the model that describes the egg seller's decision problem. He has two acts to choose between. The consequences that matter to him concern whether the omelette maker will come again to buy eggs or not. These consequences depend on the acts of the omelette maker. For example, if the omelette maker throws away the egg, he will never discover whether it is rotten or not and he will continue to buy eggs from this seller. Thus, a state of the world in the model describing the egg seller's decision problem should specify the acts of the omelette maker.

The two models that describe the decision problems of the omelette maker and the egg seller are different. In the model of each of the players the states specify the acts of the other player, but not those of the player whose decision is modeled. One may conclude that subjective probabilities of agents in interaction can be derived for each of the agents separately and game theory is not needed for Bayesian agents.

But if we analyze each player in a separate model, we miss an important aspect of the interaction, namely the reasoning of players about each other's choices. To understand this, we note that a player's beliefs are given by the probability she assigns to the various events in her space of states. If we want the player to reason about another player, and in particular about another player's beliefs, then these beliefs should be described by an event in the space of the first player. But this means that the state space of the other player should be the same as the state space of the first player. Thus, interaction of reasoning requires one model for all players. As each state in the state space of a player should include the acts of the other players, and as all the players should share the same state space, each state in this space describes the acts of all players, and hence also the consequence that results from their acts. Thus, a model that describes the interaction of actions requires a space of comprehensive states.

Comprehensive states were first studied in Aumann (1987) to facilitate the analysis of the interactive reasoning of the players in a game. Aumann claimed in this paper that the use of comprehensive states was the main novelty of his proposed model.

> The chief innovation in our model is that it does away with the dichotomy usually perceived between uncertainty about acts of nature and of personal players. [...] In our model [...] the decision taken by each decision maker is part of the description of the state of the world. (Aumann 1987)

However, in order to analyze the implications of Bayesian rationality on the players' behavior, Aumann needs each player to have a subjective probability distribution on states of the world. In this he relies on Savage's framework:

> Assume that ... as in Savage (1954), each player has a subjective probability distribution over the set of all states of the world.

But the subjective probability and the utility in Savage (1954) are derived for a state space in which neither actions nor consequences are associated with states. How can such probability and utility be derived in a comprehensive state space in Aumann (1987)?

This question is partially answered here by laying the basis for a full-fledged study of interactive decision making in a comprehensive state space. Modeling interaction of multiple agents, like Aumann (1987), requires the introduction of knowledge structures. Here we study a comprehensive state space of a single agent, which does not require the introduction of knowledge structures. The results of this research will be used in subsequent work to study the derivation of probability and utility in comprehensive state spaces of several players with knowledge structures.

1.3 Desirability and choice

It is possible to give the relation of desirability several informal intuitive meanings. We can think of one event as being more desirable than another event if learning that the first happened would make the agent happier, or more pleased, or more content, than learning that the second event happened.

Alternatively, we can think of desirability as reflecting a counterfactual choice. Although the agent does not usually have control over the events that will obtain, or at least not all of them, she can entertain the counterfactual situation in which she can choose one of two events to obtain. Saying that one event is more desirable than another means that had the agent had the opportunity to choose, she would have chosen the first event to obtain. Note that even in Savage's setup we can hardly think of the preference over acts without resorting to counterfactuals. We cannot really put the agent in situations where she can choose from *only* two acts, for any two acts. The claim that the agent prefers one act to another involves counterfactual choice: had the agent been offered only two acts from which to chose, she would have chosen this act. With this interpretation, both desirability of events and preference over acts reflect counterfactual choices.

Since both desirability and preferences over acts cannot be put to an empirical test, we need to rely on the reports of the agent about her counterfactual choices. Such reports are usually considered to be of limited reliability.

However, two recent technological innovations made closer the possibility to find out more reliably the preferences and desires of agents. New technologies of recent years suggest that reports can be validated by somatic indicators, and moreover, desirability and preferences can be found without any direct report. The first one is what we call here the biological-somatic revolution. The other is the information and communication technology (ICT) revolution.

The use of MRI, and mainly fMRI, which started about 40 years ago for brain research, was partially applied to research in decision theory. Later, and mainly in the present century, fMRI was replaced by less expensive tools like EEG, glasses to follow eye movements, etc. In many experiments the participants were asked to state their preferences between or among consumption goods. These preferences are reports of desirability. In the experiment the researcher can observe the subject's reply before the subject is aware of it. This kind of research was carried out by teams including brain researchers and researchers in departments of business schools.

ICT is an ongoing revolution that is changing its emphases. Our interest is focused on a phenomenon that has been particularly prominent in the last decade: The participation of a large majority of the population in two-way traffic on the web. (The use of smart phones exceeds that of PCs, which intensifies the participation.) Part of this traffic consists of what we termed desirability reports. The two technologies used in tandem mitigate the problem of the reliability of reports. See Telpaz and Levy (2015), Carlaw et al. (2007) and Webb et al. (2019).

1.4 An example

We illustrate the notion of desirability by the following example. Consider Eve, who is contemplating the submission of her new paper to one of several equally reputed journals between which she is indifferent. The choice of a journal is an act. There are only three consequences that matter to her: *acceptance* of the paper, *rejection*, or a request for a *revision*. Each state of the world determines the consequence of submitting the paper to each one of the said journals. In this work we assume that there are finitely many consequences. Aumann (1987) (in comment (c) of the Discussion section) suggested several reasons for finiteness.

We are now meeting Eve after making her decision to submit the paper to journal J. Now, each state is comprehensive, namely, it specifies which of the three consequences holds. In particular, the state space is partitioned into three *consequence events*: the event that consists of all states in which the paper is accepted, the event of rejection, and the event of a required revision.

Eve has a desirability relation over events and in particular over the consequence events. It is quite natural to assume that she prefers the event of acceptance over the event of revision, and the latter over the event of rejection. But the desirability relation concerns other events too. We may assume that each state of the world specifies who is the associate editor handling the paper, as this is one of the factors that determines the consequence. It is possible that Eve desires the event that Alice rather than Bob will be the associate editor handling the paper. Note that Alice handling the paper or Bob doing so, are not consequences. Eve's desire that the first event will obtain rather than the second reflects the different ways in which these two events are associated with the three consequences. For example, if it is more likely that the paper will be accepted when Alice is the associate editor, Eve may find this event more desirable than the event of Bob being the associate editor.

1.5 The axioms of desirability

We present in Sect. 2.2 seven axioms, A1–A7, on a desirability relation on a fixed comprehensive state space. Here, we sketch their gist. These axioms appear to hold for the intuitive meanings of desirability discussed above.

Before the axioms are introduced we define for any binary relation on events its *null events*. Roughly, an event is null for a given relation between two events if it does not affect the relation. More specifically, set theoretical addition (union) or subtraction of any subset of the null event to any of the two events in the relation do not change the relation between them.

Axioms A1–A3 are not special to desirability. They have analogues in other axiomatizations like Savage's, de Finneti's axioms of qualitative probability, von Neumann and Morgenstern's axioms, and many other binary relations. The first three axioms are analogous to Savage's P5, P1, and P6$'$, in this order. The *Non-degeneracy* axiom (A1) requires that the relation is non-trivial. This axiom implies that there are non-null events, which makes the next axiom of *Weak Order* (A2) non-vacuous. The latter says that the desirability relation is a complete and transitive order on the non-null events. Such axioms predate Savage and von

Neumann and Morgenstern. One of the innovative axioms of Savage that plays a crucial role in proving the existence of a probability is P6′. Our axiom A3 is similar to P6′ and we name it *Small events*, as it says that the space can be partitioned into "small" events.[2]

Axioms A4–A7 are special to desirability relations. The first among them, the axiom of *Intermediacy* (A4) says that mixed news, good and bad, lies, in terms of its desirability, between the good news and the bad news. Thus, in our example the event that the paper is either accepted or rejected is less desirable than the event of acceptance, but more desirable than the event of rejection. Formally, such an axiom is common in works that study relations on subsets.

Next, the axiom of *Persistency* (A5) is a first glimpse into the notion of likelihood, which is part of our intuition about desirability of events. Before we discuss this axiom we demonstrate how a likelihood relation between certain events can be deduced from the desirability relation. Let E and F be events that the agent equally desires, and H be an event disjoint of E and F and more desirable than both. At first glance the event $E \cup H$ cannot be more desirable than $F \cup H$. But on closer examination there can be a reason for that. If F is more likely than E, then the relative likelihood of the good news H is higher in $E \cup H$ than in $F \cup H$. Of course, likelihood is not defined in our setup, but the phenomenon just described can be used to define it. If E, F, and H are as described, and $E \cup H$ is more desirable than $F \cup H$, we say that F is more likely than E. This definition has one drawback, it depends on the event H. The axiom of persistency removes this drawback by requiring that the concept of being more likely is independent of the event H that is used to define it.

Axiom A6 of *Consequence Events* says that a consequence event is as desirable as any of its non-null subevents. Let us illustrate this axiom in our example. Consider the event 'the paper is accepted', and the more informative event 'the paper is accepted and Alice handles it'. Set-theoretically, the second event is a subevent of the first. The axiom requires that these two events are equally desirable. The reason is simple. In both events the paper is accepted. The event that Alice handles it is not a consequence and it has no value of its own, and therefore it does not change the desirability of the event that the paper is accepted.

The axioms already presented enable us to find out what the consequences are for the agent, which is impossible in Savage's theory. Savage's theory purports to derive probability and utility from observed choices. However, in order to construct the model in which this derivation takes place one needs to know in advance the consequences for the agent. But these consequences are neither observed or deduced from observations of choice. In Savage's omelette story, for example, one cannot conclude what the consequences are for the omelette maker by just observing the choices he makes about the egg. Thus, the derivation of probability and utility in Savage's model is based on one hand on the observation of choice, and on the other hand on the guesswork of consequences.

[2] Axiom P6′ is imposed on a qualitative probability relation on events. Here, it is imposed on the desirability relation on events. Axiom P6 is a translation of P6′ for a preference relation on acts.

In the model of desirability, we can find out what the consequence events are, using the data of the desirability relation. By axiom A6, a consequence event C is one that is as desirable as any more informative event. We say that such events are *complete*. It turns out that consequence events are, roughly speaking, maximal complete events. The formal details are in Sect. 2.2. Here we illustrate how in our example we can verify that the event 'Alice handles the paper' is not a consequence event, while 'the paper needs a revision' is a consequence event. When we find that the event 'Alice handles the paper and it needs a revision' is more desirable than the event 'Alice handles the paper', we conclude that the latter is not complete and therefore is not a consequence event. By contrast, we find that the event 'Alice handles the paper and it needs a revision' is as desirable as the event 'the paper needs a revision'. Moreover we find that for any event X, the event 'X and the paper needs a revision' is as desirable as the event 'the paper needs a revision'. Thus, the latter event is complete and is a candidate for being a consequence event. We only need to check that it is a maximal complete event. Indeed, suppose that A is a superevent of the event 'the paper needs a revision'. Then it contains some subevent B of either 'the paper is accepted', or 'the paper is rejected', or both. Suppose B is a subevent of 'the paper is accepted'. Then, by axiom A4 of Intermediacy, the event 'B and the paper needs a revision' is more desirable than the event 'the paper needs a revision' and hence, there are two subevents of A that are not equally desirable. We conclude that A is not a consequence event. This shows that 'the paper needs a revision' is a maximal complete event, namely a consequence event. To make this example rigorous we need to address issues concerning null events. It is easy to complete the formal description along the lines presented in Sect. 2.2.

While the axiom of Persistency (A5) enables us to define likelihood relation of equally desirable events, the axiom of Likelihood Ratio (A7) emphasizes the role of likelihood *ratios* in desirability. It amounts to saying that if the likelihood ratio of the consequence in one event is the same as in another event then the two events are equally desirable. This is done, of course, in terms of the desirability relation.

Our first theorem is:

> A desirability relation satisfies axioms A1–A7 if and only if it is represented by a probability-utility pair (P, u).

A given desirability relation can be represented by more than one probability-utility pair. Our next two theorems describe the structure of all representing pairs. We note that the probability on the state space can be given by the finitely many conditional probabilities on the consequence events, and the finite dimensional vector of the probabilities of the consequence events, which we shall call *consequence probabilities*. For two consequence probability vectors p and q, we say the p is more *optimistic* than q, if for any pair of consequences, the likelihood of the more desired one in p is higher than that likelihood in q.

Our second theorem characterizes, by two properties, a set of probabilities that can arise as the probabilities in representing pairs.

1. All the probabilities in the set have the same conditional probability on the consequence events, and thus differ only in their consequential probabilities;

2. The set of consequence probability part of the probabilities in the set is an interval, namely the convex hull of two points, and the consequence probabilities are ordered in the interval according to optimism.

We say that utility u is more *content* than v if the gains from moving to a more desirable consequence, measured by the ratio of utility difference, is higher in u than in v. In the third theorem we characterize the utilities in the set of the representing pairs.

> For any representing pair (P, u), the utility u is uniquely determined by P up to a positive affine transformation, and the probability P is uniquely determined by u. *If* (P, u) and (Q, v) are representing pairs, and the consequence probability part of P is more optimistic than the consequence part of Q, then the utility u is less content than v.

Thus the optimism in consequence probabilities is balanced by the contentment of the utility function. In our fourth theorem we show that a certain product of optimism and contentment is the same for all representing pairs.

2 The model

Let (Ω, Σ) be a *state space*, where Ω is the set of states and Σ is a σ-algebra of events. A finite set $\mathcal{C} = \{c_1, \ldots, c_n\}$ with $n \geq 2$ is the set of *consequences*. An *act* is a measurable function $f : \Omega \rightarrow \mathcal{C}$ that specifies a consequence in each state.

We fix an act f and refer to (Ω, Σ, f) as a *comprehensive state space*. This reflects the fact that each state of the world can be thought of as a *full* description of the world, including the consequence at the state specified via f.

Fixing the comprehensive space (Ω, Σ, f), we consider a binary *desirability relation*, $\succsim$, on Σ. We read $E \succsim F$ as 'E is *at least as desirable as* F'. We denote by $\sim$ the symmetric part of $\succsim$. That is, $E \sim F$ when $E \succsim F$ and $F \succsim E$. We read, $E \sim F$ as 'E is *as desirable as* F', or 'E and F are *equally desirable*'. We denote by $\succ$ the asymmetric part of $\succsim$. That is, $E \succ F$ when $E \succsim F$ but not $F \succsim E$. The relation $E \succ F$ is read as 'E is *more desirable than* F'. We introduce below the axioms A1–A7 that desirability relations should satisfy.

2.1 Null events

Given a binary relation $\succsim$ on events, we define null events as those that have no impact on the relation. In the definition that follows, we denote by $A \Delta B$, for events A and B, their symmetric difference.[3]

Definition 1 (*Null events*) An event N is *null* for the relation $\succsim$ when for all events E and F, if $E \succsim F$ $(E \not\succsim F)$, then also $E' \succsim F'$ $(E' \not\succsim F')$ for any E' and F' that satisfy $(E' \Delta E) \cup (F' \Delta F) \subseteq N$.

[3] The symmetric difference of two events consists of all the states in these events that do not belong to both, that is, $A \Delta B = (A \cup B) \setminus (A \cap B) = (A \setminus B) \cup (B \setminus A)$.

An immediate corollary of this definition is that null events do not affect any of the relations $\succ$, $\sim$, $\not\succ$, and $\not\sim$.

Corollary 1 *If E and F satisfy any of the relations $\succ$, $\sim$, $\not\succ$, and $\not\sim$, N is a null event, and $(E'\Delta E) \cup (F'\Delta F) \subseteq N$, then E' and F' satisfy the same relations.*

We denote by Σ^0 the set of null events of $\succsim$, and by Σ^+ the set of non-null events, namely, $\Sigma^+ = \Sigma \setminus \Sigma^0$. We note the following properties of Σ^0.

Claim 1 *The set of null events Σ^0 satisfies*:

1. $\emptyset \in \Sigma^0$;
2. *If N, M are in Σ^0 then $N \cup M \in \Sigma^0$*;
3. *If $N \in \Sigma^0$ and $M \subseteq N$, then $M \in \Sigma^0$.*

Proof Items 1 and 3 follow trivially from the definition of a null event.

For item 2, assume N and M are null events, $E \succsim F$, and $(E'\Delta E) \cup (F'\Delta F) \subseteq N \cup M$. We need to show that $E' \succsim F'$.

By our assumptions, there are events $N_0 \subseteq N$, $N_0' \subseteq N$ and $M_0 \subseteq M$, $M_0' \subseteq M$, such that $E \setminus E' = N_0 \cup M_0$ and $E' \setminus E = N_0' \cup M_0'$. Let $\hat{E} = (E \cap E') \cup M_0 \cup N_0'$. Then, $(E \setminus \hat{E}) \subseteq N_0$ and $(\hat{E} \setminus E) \subseteq N_0'$, and hence $E\Delta\hat{E} \subseteq N$. We analogously define $\hat{F}$ such that $F\Delta\hat{F} \subseteq N$, and since N is null, we conclude that $\hat{E} \succsim \hat{F}$. Now, $E' \setminus \hat{E} \subseteq M_0'$ and $\hat{E} \setminus E' \subseteq M_0$, thus $\hat{E}\Delta E' \subseteq M$. A similar relation holds for $\hat{F}$ and F'. Thus, as M is null, we conclude that $E' \succsim F'$. The proof for the case that $E \not\succsim F$ is similar. □

The three properties of Σ^0 in Claim 1 make Σ^0 an *ideal* of events in Σ. Savage (1954) also proves that the null events defined in his model form an ideal. Finally, the set of null events of a probability measure is obviously an ideal.

Without making any assumption about $\succsim$, it may happen that all events are null. However, we next show that if this relation is non-trivial, then there necessarily exist non-null events.

Claim 2 *If there are events E and F such that $E \succ F$, then there are non-null events.*

Proof Assume that $E \succ F$ and suppose that, contrary to the claim, all events are null. Set $E' = F$ and $F' = E$. Then $E'\Delta E$ and $F'\Delta F$ are null, and thus $E' \succsim F'$, that is, $F \succsim E$. Thus, $E \sim F$, a contradiction. □

Finally, it is easy to prove

Claim 3 *If Σ^0 is the set of null events of $\succsim$, and $\succsim'$ is the restriction of $\succsim$ to the non-null events of $\succsim$, that is to $(\Sigma^+)^2$, then Σ^0 is also the set of null events of $\succsim'$.*

2.2 The axioms of desirability

The first three axioms are typical of many binary relations and do not reflect the intuitive meaning of desirability. The first axiom requires that the desirability relation is non-degenerate. It is a mild assumption, since without it there is nothing of interest to say about the given relation. Non-degeneracy is assumed also in Savage (1954), as well as in the axioms of qualitative probability in de Finetti (1931).

A1 (*Non-degeneracy*) There are events E and F such that $E \succ F$.

Axiom A1 of Non-degeneracy guarantees, by Claim 2, that there are non-null events. We are interested in the desirability relation only between the non-null events of $\succsim$. In the next axiom we require that $\succsim$ be defined only on pairs of non-null events, and on these events it is a weak order, namely complete and transitive.

A2 (*Weak Order*) $\succsim$ is contained in $(\Sigma^+)^2$ and it is a complete and transitive relation.

Alternatively, we could allow $\succsim$ to be defined also out of $(\Sigma^+)^2$ and require only that $\succsim'$, the restriction of $\succsim$ to $(\Sigma^+)^2$, is a weak order on it. By Claim 3, $\succsim$ and $\succsim'$ have the same set of non-null sets. Since we are going to compare the desirability of non-null events only, we can use either $\succsim$ or $\succsim'$, and the latter satisfies axiom A2.

Next, we require that, given a strict desirability relation between two events, the state space can be partitioned into events that are small in the sense that they do not affect the given relation. This axiom is a slight variation of property P6$'$ in Savage (1954).

A3 (*Small Events*) For two events E and F such that $E \succ F$ there exists a partition of Ω, $\Pi = (\Pi_1, \ldots, \Pi_m)$, such that for each i, if $F' \Delta F \subseteq \Pi_i$, then $E \succ F'$, and if $E' \Delta E \subseteq \Pi_i$, then $E' \succ F$.

The next four axioms capture the intuitive meaning of desirability. We first formalize the idea that a mixture of some good news and some bad news is more desirable than the bad news and less desirable than the good news. It has the same spirit as the averaging condition in Bolker (1967). We illustrate it with the example discussed in the introduction. Let E be the event *Alice handles the paper* and F, which is disjoint from event E, that *Bob handles the paper*. Suppose that E is weakly more desirable than F, that is $E \succsim F$. The event $E \cup F$ is mixed news. Therefore E, the good news, must be at least as desirable as $E \cup F$, and $E \cup F$ must be at least as desirable as the less desirable event F.

A4 (*Intermediacy*) Let E and F be disjoint non-null events. Then the relations $E \succsim F$, $E \cup F \succsim F$, and $E \succsim E \cup F$ are equivalent.

We illustrate the next axiom using our example. In doing so we use axiom A6 of Consequence Events which is presented below, but was already discussed in the introduction. Consider the events $A =$ *acceptance* and $B =$ *acceptance and Alice handles the paper*, which by axiom A6 of Consequence Events are equally desirable as only acceptance is a consequence about which Eve cares. Note, however, that as $B \subseteq A$, B is less likely than A. Now consider the event $G =$ *revision*, which is disjoint from A and B, and the events $A \cup G$ and $B \cup G$. The likelihood of *acceptance* in $A \cup G$ is higher than in $B \cup G$. Therefore, $A \cup G$ should be more desirable than $B \cup G$. Note that if G is an event disjoint from A and B that is less desirable than both A and B, then we expect that the relation of desirability between $A \cup G$ and $B \cup G$ will be reversed.

The implication in this example is based on an informal, intuitive notion of likelihood relation. But we can reverse the reasoning and use this to formally define a restricted concept of likelihood relation.

Definition 2 Suppose that $A \sim B$ and that G is a non-null event such that $G \cap (A \cup B) = \emptyset$. Then *A is at least as likely as B according to G* if either $A \succ G$ and $A \cup G \succsim B \cup G$, or $G \succ A$ and $B \cup G \succsim A \cup G$.

If the relation of likelihood according to G is to capture the likelihood of equally desirable events, then this relation should not depend on G. That is, if we take instead of G another event H with the same properties, then the relation of likelihood according to H should be the same. This is the meaning of the next axiom, which is in the spirit of the impartiality property in Bolker (1967).

A5 (*Persistency*) Suppose $A \sim B$, and G, H are non-null events disjoint of A and B such that $G \not\sim A$ and $H \not\sim A$. If A is at least as likely as B according to G, then A is also at least as likely as B according to H.

By axiom A5 of Persistency, we can now define the following relation between events.

Definition 3 For two equally desirable events A and B, A is *at least as likely as B* if A is at least as likely as B according to *some* G.

The first five axioms did not involve consequences. The next two axioms address consequences directly. This is where our model deviates from Jeffrey (1965). We first introduce some notation. For $c \in \mathcal{C}$ we denote by C the event that the consequence of f is c. Namely, $C = \{\omega \mid f(\omega) = c\}$. We call the events C *consequence events*. For each E and c we write E_c for $E \cap C$. Thus, the event E is the disjoint union of the events E_c for all consequences c.

The next axiom addresses the nature of consequence events that distinguishes them from other events. Such distinction does not exist in Savage's setup, as consequence events do not exist. The axiom says that when the agent is informed that a consequence c occurs, any additional information is irrelevant to desirability. Formally:

A6 (*Consequence Events*) For any consequence c and a non-null event $E \subseteq C$, $E \sim C$.

Axiom A6 of Consequence Events and axiom A4 of Intermediacy enable us to *derive* consequences from the desirability relation, rather than guess them as in Savage's model. To see this, let's say that an event E is *complete* if all its non-null subevents are equally desirable (that is, the restriction of $\succsim$ to the non-null events of E is a complete order). Then, axiom A6 states that a consequence event is complete. Hence, only the complete events are candidates for being consequence events. Obviously not all of them are, as every subevent of a complete event is itself a complete event and in particular, subevents of a consequence event are complete events. However, the next claim helps us to discover consequence events.

Claim 4 *Suppose that any two non-null consequence events are not equally desirable. Then, a non-null consequence event C is maximally complete in the sense that if $C \subseteq A$ and A is complete, then $A \setminus C$ is null.*

Indeed, suppose that $C \subseteq A$ and $A \setminus C$ is not null. Then there exists a non-null consequence event $D \neq C$ and a non-null event $D' \subseteq D$, such that $C \subset C \cup D' \subseteq A$. Since, by axiom A6, $D' \sim D$, axiom A4 of Intermediacy implies that if $D \succ C$ then $C \cup D' \succ C$, and if $C \succ D$ then $C \succ C \cup D'$. In either case, this shows that A is not complete.

The family of maximal complete events includes not only consequence events, but also any event that differs from a consequence event by a null event, that is, any M such that $M \Delta C$ is null for a non-null consequence event C. Thus, we are able to identify consequence events up to null events. In case there are equally desirable consequence events, say C and D, then by axiom A4, $C \cup D$ is an complete and thus C and D are not maximal. However the union of consequence events that are as desirable as C is a maximally complete, and for the purpose of finding out the probability and utility of the agnet it is enough to consider this union as a consequence event.

The next axiom, like axiom A5 of Persistency, addresses issues of likelihood. Axiom A5 gave rise to Definition 3 which introduced the likelihood relation on equally desirable events. The next axiom uses a qualitative proxy to likelihood *ratios*. Thus, it is a relation between two pairs of events, more specifically, pairs (E_c, E_d) and (F_c, F_d) for two distinct consequences c and d. If C is at least as desirable as D, then the relation $E_c \cup E_d \succsim F_c \cup F_d$ can be interpreted as a qualitative expression that is represented by a quantitative relation of a higher likelihood ratio. Namely, the numerical likelihood ratio of E_c to E_d is at least as high as the likelihood ratio of F_c to F_d. Thus, to say that likelihood ratio of the two pairs is the same amounts to saying that $E_c \cup E_d \sim F_c \cup F_d$. We can now say that desirability depends on the likelihood ratio of consequences by saying that if the likelihood ratios of consequences in E are the same as in F then E and F are equally desirable. To say that the likelihood ratio of consequences is the same in E and F is to say that for each c and d, $E_c \cup E_d \sim F_c \cup F_d$. Of course, we need to be careful to state that events are equally desirable only when they are non-null, which we do formally next.

A7 (*Likelihood ratio*) Let E and F be non-null events. If for each pair of distinct consequences, c and d, $E_c \cup E_d$ and $F_c \cup F_d$ are either both null or both non-null, and in the latter case $E_c \cup E_d \sim F_c \cup F_d$, then $E \sim F$.

3 The main theorems

Our first result concerns representation of a desirability relation in a comprehensive state space. For this we define how a probability-utility pair represents a desirability relation.

Definition 4 (*Representation*) Consider a pair (P, u), where P is a finitely additive probability on (Ω, Σ, f) and $u : \mathcal{C} \to \mathbb{R}$. We say that (P, u) *represents* a binary relation $\succsim$ on Σ if the set of null events of $\succsim$ is the set of P-null events, and for all non-null events A and B, $A \succsim B$ if and only if

$$\sum_{i=1}^{n} u(c_i)P(C_i|A) \geq \sum_{i=1}^{n} u(c_i)P(C_i|B). \tag{1}$$

Note that if inequality (1) holds, then it holds also for any positive affine transformation of u, i.e., $u \mapsto \alpha u + \beta$ where $\alpha > 0$.

Theorem 1 *For a comprehensive state space (Ω, Σ, f), a relation $\succsim$ on Σ satisfies axioms $A1$–$A7$ if and only if there exists a pair $(P,\ u)$ that represents it. Moreover, P is non-atomic.*[4]

We illustrate the relation between probability-utility pairs and desirability relations in the following example.

Example 1 Let the state space (Ω, Σ) be the unit interval with the σ-algebra of Borel sets. The set of consequences is $\mathcal{C} = \{c_1, c_2, c_3\}$. The fixed act f is defined by $f(\omega) = c_1$ for $\omega \in [0, 1/3)$, $f(\omega) = c_2$ for $\omega \in [1/3, 2/3)$, and $f(\omega) = c_3$ for $\omega \in [2/3, 1]$. Thus, the consequence events are: $C_1 = [0, 1/3)$, $C_2 = [1/3, 2/3)$, and $C_3 = [2/3, 1]$. The comprehensive state space is (Ω, Σ, f).

Consider the pair $(P,\ u)$, where P is the uniform probability distribution, and the utility function, $u : \mathcal{C} \to \mathbb{R}$, is given by $u(c_1) = u_1 = 0$, $u(c_2) = u_2 = 1/2$, and $u(c_3) = u_3 = 1$. Denote by P_i, for $i = 1, 2, 3$, the conditional probability of P on C_i. For a P-non-null event E, let $x_i = P(E|C_i)$. Then the conditional utility given E is:

$$[(0)(1/3)x_1 + (1/2)(1/3)x_2 + (1)(1/3)x_3]/[(1/3)x_1 + (1/3)x_2 + (1/3)x_3].$$

The conditional expectation defines a desirability relation $\succsim$ on the P-non-null events, which $(P,\ u)$ represents. The null events of $\succsim$ are the P-null-events. Note that the conditional expectation given E is determined by the x_i's. Thus, in particular, if two events have the same conditional probability given each C_i, then they are equally desirable.

In order to simplify the formulation of the following results we make two assumptions.

Assumptions

1. For each consequence c, the event C is non-null,
2. $C_n \succ C_{n-1} \succ \cdots \succ C_1$.

The main thrust of the second assumption is that no two distinct events C_i and C_j are similar. The ordering of desirability according to the indices is made, of course, without loss of generality.

[4] A probability measure P is non-atomic if for each event E and α in [0, 1], there exists an event $F \subseteq E$ such that $P(F) = \alpha P(E)$. This condition appeared first in Savage (1954) and was described as non-atomicity by Machina and Schmeidler (1992). Gilboa (1987) defined an extension of this condition to non-additive measures, and referred to a measure that satisfies it as convex ranged.

The question that usually arises in representation theorems is the uniqueness of presentation. In our case the set of pairs that represent $\succsim$ is not a singleton. In the following theorems we characterize this set. We denote by $\mathcal{P}(\succsim)$ the set of all probability measures P such that for some u, (P, u) represents $\succsim$.

We decompose a probability P on (Ω, Σ) into two parts: The *conditional* part $(P_i)_{i=1}^n$, where for each i, $P_i(\cdot) = P(\cdot \mid C_i)$, and the *consequential* part, p, in the simplex $\Delta(\mathcal{C})$, where $p_i = P(C_i)$. Thus, for each event E, $P(E) = \sum_{i=1}^n p_i P_i(E)$. It turns out that the conditional part is uniquely determined for the given desirability relation, while the consequential part is not. In order to describe this non-uniqueness we introduce the notion of *optimism*.

For two positive probabilities p and q in $\Delta(\mathcal{C})$, we say that p is *more optimistic* than q, and write $p \gg q$, if for each $i<j$, $p_j/p_i > q_j/q_i$. The reason why these inequalities describe optimism follows from Assumption 2. If $p \gg q$, then for any two consequences the likelihood of the preferred one is higher in p than in q. Let $\rho(p)$ be the $(n-1)$-dimensional vector with components $\rho_i(p) = p_{i+1}/p_i$ for $i = 1, \ldots n-1$. We say that p *likelihood-ratio dominates*[5] q if $\rho(p) > \rho(q)$. Obviously, p is more optimistic than q if and only if p likelihood-ratio dominates q.

An open interval of positive probabilities $(p, q) = \{\alpha p + (1-\alpha)q \mid 0<\alpha<1\}$ is *ordered by optimism* if for any $\alpha > \alpha'$, $\alpha p + (1-\alpha)q \gg \alpha' p + (1-\alpha')q$. The interval is *maximal* if p and q are on the boundary of the simplex $\Delta(\mathcal{C})$.

We are now ready to describe the multiplicity of the probabilities in the representing pairs.

Theorem 2 *A set of probabilities $\mathcal{P}$ is $\mathcal{P}(\succsim)$ for some relation $\succsim$ on Σ that satisfies axioms A1–A7 and Assumptions 1,2 if and only if:*

1. *The conditional parts of the probabilities in $\mathcal{P}$ coincide. That is, for each P and Q in $\mathcal{P}$, $(P_i) = (Q_i)$,*
2. *The consequential parts of probabilities in $\mathcal{P}$ form a maximal interval ordered by optimism.*

Finally, we characterize the utilities in the representing pairs.

Theorem 3 *For every $P \in \mathcal{P}(\succsim)$, there exists a utility u, which is unique up to a positive affine transformation, such that (P, u) represents $\succsim$. Moreover, if (P, u) and (Q, u) represent $\succsim$, then $P = Q$.*

We can say more about the representing utilities. Let $u_i = u(c_i)$ and define the vector of *utility gains* $\Delta u = (\Delta u_i)_{i=1}^{n-1}$ by $\Delta u_i = u_{i+1} - u_i$. By Theorem 1*, $\Delta u > 0$. For two utility vectors u and v we say that u is *more content* than v if for each $i<j$ between 2 and n, $\Delta u_j/\Delta u_i < \Delta v_j/\Delta v_i$. The $(n-2)$- dimensional vector $\rho(\Delta u)$, where $\rho_i(\Delta u) = \Delta u_{i+1}/\Delta u_i$ for $i = 2, \ldots, n-1$ is the vector of the *utility-gain ratio*. Obviously, u is more content than v if and only if $\rho(\Delta v) > \rho(\Delta u)$, that is, $\Delta v \gg \Delta u$. Note that $\rho(u)$ is invariant under positive affine transformations of u.

[5] It is straightforward to see that Likelihood-ratio dominance implies stochastic dominance.

Roughly speaking, being more optimistic means assigning higher probability to more desirable consequences, and being more content means having less utility from such consequences. The next theorem says that being more optimistic is balanced by being more content.

Theorem 4 *For each $i = 2, \ldots, n-1$, the product $\rho_i(\Delta u)\rho_i(p)\rho_{i-1}(p)$ is the same for all pairs (P, u) that represent $\succsim$. Thus, if (P, u) and (Q, v) represent $\succsim$, and $\rho(p) > \rho(q)$, then $\rho(\Delta u) < \rho(\Delta v)$.*

Example 2 The desirability relation in Example 1 can be represented by other pairs (Q, v). By Theorem 2, the conditional probability of Q given C_i is P_i for each i. Thus, $Q = q_1P_1 + q_2P_2 + q_3P_3$ for some probability vector $q = (q_1, q_2, q_3)$. If we choose $q = (1/6, 1/3, 1/2)$ and $v_1 = 0, v_2 = 3/4$ and $v_3 = 1$, then (Q, v) also represents $\succsim$. This can be easily seen by checking that the conditional expected utilities of the two pairs are similarly ordered. Note that $\rho(q) = (2, 3/2)$, while for $p = (1/3, 1/3, 1/3)$, in Example 1, $\rho(p) = (1, 1)$. Thus, $\rho(q) > \rho(p)$, and therefore q is more optimistic than p. Also $\Delta u = (1/2, 1/2)$, $\Delta v = (3/4, 1/4)$, and hence $\rho(\Delta u) = (1)$ and $\rho(\Delta v) = (1/3)$, which demonstrates the claim of Theorem 4.

In Section 5.10, we show how to compute the maximal interval of probability vectors that are ordered by optimism guaranteed by Theorem 2.

4 Literature survey

Jeffrey (1965) introduced a real-valued function on *propositions* which he called Desirability. The set of propositions was rigorously modeled by Bolker (1967) as a complete Boolean algebra. Measures are defined on such algebras in much the same as on fields or sigma fields, which are, in particular, Boolean algebras of events. The desirability function in Jeffrey (1965) and Bolker (1967) is the ratio of a signed measure and a probability measure on the Boolean algebra.

The theory of desirability presented by Jeffrey and Bolker is a departure from the theories of von Neumann and Morgenstern (1953) and Savage (1954) in that it does not include consequences (or prizes) and a utility function on consequences. Acts cannot be defined in their theory since there are no consequences. Bolker comments on the difference between Jeffrey's model and Savage's model: "The states must be unambiguously described. By so doing we blur the often useful distinctions among acts, consequences and events" (Bolker 1967, foonote 7). This lost distinction is reinstated here, where we use Savage's model in which consequences are the main features.

Based on a previous work, Bolker (1966, 1967) considered a binary relation on propositions, which was not named, and axioms on this relation that guarantee that it can be represented by a desirability measure. He has two axioms that correspond to our axioms A4 of Intermediacy and A5 of Persistency.

In addition to the essential difference of incorporating consequences, our model differs from Bolker's model in other aspects. (1) In Bolker (1967) the relation is defined on the non-zero elements of a complete non-atomic Boolean algebra. This corresponds to the quotient space of a measurable space with respect to null events.

Thus, null events must be defined prior to the definition of the desirability relation. In our model, as in Savage's, null events are defined in terms of the relation rather than assumed. (2) Bolker assumes that the relation is continuous and derives representing probabilities that are σ-additive. We make no continuity assumption and, like Savage (1954) and unlike Bolker (1967), we derive a finitely additive probability.

Bolker (1967) and Jeffrey (1983) describe a linear structure of the set of probability-utility pairs that represent the binary relation in their model, a structure that was suggested to Jeffrey by Kurt Gödel (Jeffrey 1983, p. 143). The characterization of this set in our model depends on its central feature, the set of consequences. We show: (1) the conditional probability given a consequence event is uniquely determined; (2) the probabilities of the consequence events are ordered by optimism; and (3) a cardinal utility for a given probability is uniquely determined and the utility gains are ordered by contentment.

Binary relations on subsets of a given set were studied in various works. de Finetti (1931) considered a relation on events in a state space, named *qualitative probability*. He proposed several axioms on qualitative probability, but they were not enough to guarantee that qualitative probability can be represented by a numerical probability. By adding axiom P6′, Savage (1954) showed that a qualitative probability has a unique representation by an additive probability. We use this result to prove the existence of a unique probability on each consequence event, using axiom A3 of Small Events, which is similar to Savage's P6′.

Ahn (2008) studied a preference over lotteries. His axioms, like ours, resemble those of Bolker, despite the different domain of the relation. Thus, his presentation of the relation is expressed as a ratio of an integral on utility divided by a probability.

Luce and Krantz (1971) used conditional expected utility to represent a binary relation. However, unlike desirability, the relation they study is not defined on events, but on *conditional acts*, namely acts that are a function not on the whole state space, but only on an event in this space. For a further discussion of the relation between Luce and Krantz (1971), Savage (1954), and Jeffrey (1965), see Chai et al. (2016).

5 Proofs

5.1 An outline of the proofs

We omit the proof of the simple "if" part of Theorem 1. We prove first a restricted version of Theorem 1 under Assumptions 1 and 2, and then show how Theorem 1 can be derived from this version.

Theorem 1* *For a comprehensive state space* (Ω, Σ, f), *a relation* $\succsim$ *on* Σ *satisfies axioms* A1–A7 *and Assumptions 1 and 2 if and only if there exists a pair* (P, u) *that represents* $\succsim$, *such that* $P(C_i) > 0$ *for* $i = 1, \ldots, n$ *and* $u(c_n) > u(c_{n-1}) > \cdots > u(c_1)$.

In Sect. 5.2 we derive for each consequence c a probability P_c on Σ_c, the σ-field of events in C, that will serve as the conditional probability of the probability P in Theorem 1*. Definition 2 enables us to introduce a likelihood relation on a family of similar events. Since, by axiom A6 of Consequence Events, all non-null subevents of C are similar, we manage to define a likelihood relation on Σ_c. This relation is shown to be a qualitative probability. By axiom A3 of Small Events, a theorem of Savage ensures that there exists a unique non-atomic probability on Σ_c, which represents the qualitative likelihood relation on Σ_c.

In Sect. 5.3 we show that the desirability relation between events depends only on the n-dimensional vector of their conditional probabilities $(P_c(E \cap C))_{c \in \mathcal{C}}$. Moreover, it is homogeneous in this vector.

This enables us to translate, in Sect. 5.4, the desirability relation on events into a relation on the positive orthant of $\mathbb{R}^{\mathcal{C}}$. We show that the sets defined by this latter relation are convex, and describe their topological properties.

In Sect. 5.5 we again use Definition 2 to define a relation of being more likely on each equivalence class of points in $\mathbb{R}^{\mathcal{C}}$. We characterize the convexity of sets defined in terms of this relation and their topological properties.

We show in Sect. 5.6 that the sets of being more likely than x and less likely than x, in the set of points equivalent to x, can be separated by a probability vector. Moreover, this vector is independent of x. Such a probability vector will be the probability of the consequence events.

In Sect. 5.7 we characterize the space of separating functionals of the previous subsection in terms of exchange rates of coordinates in the Euclidean space. These exchange rates help us to derive the utility in the next subsection.

Using the conditional utility in Sect. 5.2, the probabilities derived in Sect. 5.5, and the utility derived in Sect. 5.8, we go back to the desirability relation and prove Theorems 1–4.

5.2 The conditional probability over consequences

The following are three immediate corollaries of axioms A4 of Intermediacy and A2 of Weak Order. The first is not only a corollary of the two axioms, but combined with axiom A2 implies axiom A4.

Corollary 2 *If E and F are disjoint non-null events, then the relations $E \succ F$, $E \cup F \succ F$, and $E \succ E \cup F$ are equivalent.*

Corollary 3 *If E and F are disjoint non-null events, then the relations $E \sim F$ and $E \cup F \sim F$ are equivalent. Hence, if $E^1, \ldots, E^k$ are non-null events that are disjoint in pairs, and $E^1 \sim E^2 \sim \ldots \sim E^k$, then $\cup_{i=1}^k E^i \sim E_1$.*

Corollary 4 *Let E and F be disjoint events. If $A \succ E$ and $A \succsim F$, then $A \succ E \cup F$. If $E \succ A$ and $F \succsim A$, then $E \cup F \succ A$.*

Proof For the first part, if $E \succsim F$, then by intermediacy $A \succ E \succsim E \cup F$. If $F \succ E$, then by Corollary 1, $A \succsim F \succ E \cup F$. The second part is similarly proved. $\square$

We denote by Σ_c the σ-algebra that Σ induces on C, namely, $\Sigma_c = \{E \mid E \subseteq C, E \in \Sigma\}$.

We begin with a derivation of a non-atomic probability distribution P_c on Σ_c for each consequence c. This is done by defining a relation $\gtrapprox$ on Σ_c, in terms of the relation $\succsim$, and showing that it satisfies the axioms of qualitative probability.

Fix for now a consequence c and the corresponding event C. Choose a non-null event G such that $G \cap C = \emptyset$ and $G \not\sim C$. By Assumption 2, and since $n \geq 2$, there exists such a G, as C_j for $j \neq i$ satisfies it. Note, that since G is non-null, for any $A \in \Sigma_c$, including the null events, $A \cup G$ is non-null. We define a binary relation $\gtrapprox$ on Σ_c as follows.

Definition 5 For $A, B \in \Sigma_c$, $A \gtrapprox B$ if either $C \succ G$ and $A \cup G \succsim B \cup G$, or $G \succ C$ and $B \cup G \succsim A \cup G$.

Observe that non-null events A and B in Σ_c are similar events by axiom A6 of Consequence Events, and therefore $A \gtrapprox B$ if and only if A is more likely than B according to G, as in Definition 2. Thus, $\gtrapprox$ is an extension of the latter relation to all events in Σ_c. We write $A \approx B$ when $A \gtrapprox B$ and $B \gtrapprox A$, and $A > B$ when it is not the case that $B \gtrapprox A$.

Proposition 1 *There exists a unique probability measure P_c on Σ_c such that for any $A, B \in \Sigma_c$, $A \gtrapprox B$ if and only if $P_c(A) \geq P_c(B)$. The probability P_c is non-atomic.*

Proof We first show that $\gtrapprox$ is a qualitative probability on Σ_c. That is, it satisfies the following properties for all A, A', and B in Σ_c such that $B \cap (A \cup A') = \emptyset$.

1. $\gtrapprox$ is transitive and complete;
2. $A \gtrapprox A'$ if and only if $A \cup B \gtrapprox A' \cup B$;
3. $A \gtrapprox \emptyset$, $C > \emptyset$.

Since $G \not\sim C$, either $G \succ C$ or $C \succ G$. We assume that $C \succ G$. The proof for the other case is analogous.

By Weak Order either $A \cup G \succsim B \cup G$, in which case $A \gtrapprox B$, or $B \cup G \succsim A \cup G$, in which case $B \gtrapprox A$. Thus, $\gtrapprox$ is complete. Suppose that $A_1 \gtrapprox A_2$ and $A_2 \gtrapprox A_3$. Then, $A_1 \cup G \succsim A_2 \cup G$ and $A_2 \cup G \succsim A_3 \cup G$. By Weak Order $A_1 \cup G \succsim A_3 \cup G$, and thus $A_1 \gtrapprox A_3$. Therefore $\gtrapprox$ is transitive.

To show 2, we consider the following four cases. (a) B is null. In this case, $A \cup B \cup G \succsim A' \cup B \cup G$ if and only if $A \cup G \succsim A' \cup G$, which yields 2. (b) A is null and A' is not. This case is impossible when $A \gtrapprox A'$, because by Corollary 2, $A' \cup G \succ G \sim A \cup G$. (c) A is non-null and A' is null. By Intermediacy $A \cup G \succsim G \sim A' \cup G$. Thus, in this case, necessarily $A \gtrapprox A'$. Since $B \succ G$, $A \sim B \succsim B \cup G$, and hence by axiom A4 of Intermediacy, $A \cup B \cup G \succsim B \cup G \sim A' \cup B \cup G$. Thus, in this case it is also necessary that $A \cup B \gtrapprox A' \cup B$. (d) All three events A, A' and B are non-null. In this case, $A \gtrapprox A'$ means that A is more likely than A' according to G. As $B \sim C \succ G$, it follows by Corollary that $C \sim B \succ B \cup G$. Also $(B \cup G) \cap (A \cup A') = \emptyset$. Thus, by axiom A5 of Persistency, A is more likely than A' according to G if and only if A is more likely than A' according to $B \cup G$. Hence, $A \cup G \succsim A' \cup G$ if and only if $A \cup B \cup G \succsim A' \cup B \cup G$.

If A is non-null, then by Corollary 2, $A \cup G \succ G = \emptyset \cup G$. Hence it is not the case that $\emptyset \cup G \succsim A \cup G$, and therefore $A > \emptyset$. In particular, $C > \emptyset$. If A is null then $A \cup G \sim \emptyset \cup G$. Which shows that for all A, $A \gtrapprox \emptyset$. This proves 3.

Next, we prove a property of $\gtrapprox$ which is named by Savage P6':

If $E > F$, then there exists a finite partition of C, $(\Pi_i)_{i=1}^{k}$, such that for each i, $E > F \cup \Pi_i$.

Since $E > F$, it follows that $E \cup G \succ E \cup G$. Let $\{\Pi'_i \mid i = 1, \ldots, m\}$ be the partition the existence of which is guaranteed by axiom A3 of Small Events for the last relation. Then, the set of nonempty events of the form $\Pi_i = \Pi'_i \cap C$ is a partition of C and for each such event Π_i, $(F \cup G \cup \Pi_i)\Delta(F \cup G) = P_i \subseteq \Pi'_i$. Thus, by the said axiom, $E \cup G \succsim F \cup G \cup \Pi_i$, which means $E > F \cup \Pi_i$.

This property with the properties of $\gtrapprox$ as qualitative probability imply the claim of the proposition as is shown in Savage (1954). □

In the next subsection we show that the desirability of an event E depends only on the probabilities $P_c(E_c)$. Here, we show that the question whether E is null or not depends only on these probabilities.

Definition 6 Let $\pi : \Sigma \to \mathbb{R}^{\mathcal{C}}$ be defined by $\pi(E) = (P_c(E_c))_{c \in \mathcal{C}}$.

Proposition 2 *An event N is null if and only $\pi(N) = 0$.*

Proof Since Σ^0 is closed under unions, and inclusion, an event N is null if and only if for each c, N_c is null. Thus, it is enough to show that N_c is null if and only if $P_c(N_c) = 0$. If N_c is null then for any non-null H, $N_c \cup H \sim H$ and therefore $N_c \approx \emptyset$ and thus, $P_c(N_c) = 0$. For the converse suppose $P_c(N_c) = 0$. We need to show that if $E \succsim F$, $E\Delta E' \subseteq N_c$, and $F\Delta F' \subseteq N_c$ then $E' \succsim F'$. For this it suffices to show that $E \sim E'$ and $F \sim F'$. Note that $E \setminus C = E' \setminus C$. Now, if $E \setminus C \sim C$, then by Corollary $E = E_c \cup (E \setminus C) \sim C$ and similarly $E' \sim C$ and we are done. Otherwise, $E \setminus C \not\sim C$. Now, $E_c = (E_c \cap E'_c) \cup N_c^1$ for some $N_c^1 \subseteq N_c$. Since $P_c(N_c^1) = 0$, it follows by axiom A5 of Persistency, that $E = (E_c \cap E'_c) \cup N_c^1 \cup (E \setminus C) \sim (E_c \cap E'_c) \cup (E \setminus C)$. Similarly $E' \sim (E_c \cap E'_c) \cup (E' \setminus C)$. Since $E \setminus C = E' \setminus C$, it follows that $E \sim E'$. Similarly, $F \sim F'$. □

5.3 The homogeneity of desirability

In this subsection we prove:

Proposition 3 *If there exists $t > 0$ such that $\pi(E) = t\pi(F) \neq 0$, then $E \sim F$.*

To prove it we use the following three lemmas.

For each non-null G, the *support* of G is $\mathcal{C}(G) = \{c \mid G_c \text{ is non-null}\}$. We split the support into two parts $C^-(G) = \{c \in \mathcal{C}(G) \mid G \succ G_c\}$ and $\mathcal{C}^+(G) = \{c \in \mathcal{C}(G) \mid G_c \succsim G\}$.

Lemma 1 *The set $C^+(G)$ is not empty, and if $|\mathcal{C}(G)| \geq 2$, then also $\mathcal{C}^-(G)$ is not empty.*

Proof Suppose that $C^+(G) = \varnothing$. Then $G = \cup_{c \in C^-(G)} G_c$. By Corollary 4, $G \succ \cup_{c \in C^-(G)} G_c$, which is impossible. Assume now that $|\mathcal{C}(G)| \geq 2$ and suppose that $C^-(G) = \varnothing$. Then for some c and d in $C^+(G)$, $G_c \succ G_d$. Again by Corollary 4, $G = \cup_{c \in C^+(G)} G_c \succ G$. $\square$

Lemma 2 *Let G be an event such that $|\mathcal{C}(G)| \geq 2$. Denote for each event X such that $\mathcal{C}(X) = \mathcal{C}(G)$, $X^+ = \cup_{c \in \mathcal{C}^+(G)} X_c$ and $X^- = \cup_{c \in \mathcal{C}^-(G)} X_c$. If $G^+ \subset X^+$ and $X^- \subset G^-$, and the events $G^+ \setminus X^+$ and $X^- \setminus G^-$ are non-null, then $X \succ G$.*

Proof By Corollary 4, $G \succ G^- \setminus X^-$. This implies that $X^- \cup G^+ \succ G$, because if $G \succsim X^- \cup G^+$, then $G \succ (X^- \cup G^+) \cup (G^- \setminus X^-) = G$. Also, $X^+ \setminus G^+ \succsim G$. Hence, $(X^- \cup G^+) \cup (X^+ \setminus G^+) \succ G$. Since $\mathcal{C}(X) = \mathcal{C}(G)$ it follows that $X \sim (X^- \cup G^+) \cup (X^+ \setminus G^+)$ and thus $X \succ G$. $\square$

Next, we describe a simple result of axiom A3 of Small Events. If $F \succ E$, and $E^1 \subseteq E$ is non-null, then there exists $D \subseteq E^1$ such that $D \cap E^1$ is non-null and $F \succ E \setminus D$. Indeed, choose the partition Π in axiom A3, and select an element Π_i of Π such that $\Pi_i \cap E^1$ is non-null, and set $D = \Pi_i \cap E^1$. This result can be generalized as follows.

Lemma 3 *If $F \succ E$, and $E^1, \ldots, E^m$ are non-null subevents of E, then there exists $D \subseteq \cup_{i=1}^m E^i$ such that for each i, $D \cap E^i$ is non-null and $F \succ E \setminus D$.*

Proof Prove by induction on m. In the k stage we have D^k that satisfies the condition for $E^1, \ldots, E^k$. Since $F \succ E \setminus D^k$, we can apply axiom A3 of Small Events and choose P_i such that $P_i \cap E^{k+1}$ is non-null. We let $D^{k+1} = (D^k \cup P_i) \cap \cup_{i=1}^{k+1} E^i$. $\square$

Proof of Proposition 3 By Proposition 2, $\mathcal{C}(E) = \mathcal{C}(F) = \{c \mid p_c(E_c) = p_c(F_c) > 0\}$. If this set, which we denote by $\mathcal{C}$, is a singleton c, then both E and F are similar to C and we are done. We assume therefore that $|\mathcal{C}| \geq 2$.

We prove first for $t = 1$. By the definition of p_c and axiom A5 of Persistency, for each $d \neq c$ in $\mathcal{C}$, $E_c \cup F_d \sim F_c \cup F_d$. Similarly, by the definition of p_d, $E_c \cup F_d \sim E_c \cup E_d$. Thus, $E_c \cup E_d \sim F_c \cup F_d$. It follows by axiom of Likelihood Ratio that $E \sim F$.

Suppose that $t = k/m$ for some integers k and m. By the non-atomicity of p_c, there exists for each $c \in \mathcal{C}$, a partition $E_c^1, \ldots, E_c^k$ of E_c into k equally p_c-probable events and a partition $F_c^1, \ldots, F_c^m$ of F_c into m equally p_c-probable events. Then $p_c(E_c^i) = p_c(F_c^j)$ for all $c \in \mathcal{C}$ and i, j. Let $E^i = \cup_{c \in \mathcal{C}} E_c^i$ and $F^j = \cup_{c \in \mathcal{C}} F_c^j$. Then, by the claim for $t = 1$, $E^i \sim F^j$ for all i and j. As all the E^i's are disjoint in pairs and similar, it follows by Corollary 3 that $\cup_{i=1}^k E^i \sim E^1$. In the same way, $\cup_{j=1}^m F^j \sim F^1$. Since for all $c \notin \mathcal{C}$, E_c and F_c are null, $E \sim \cup_{i=1}^k E^i$ and $F \sim \cup_{j=1}^m F^j$. But, $E^1 \sim F^1$, and therefore $E \sim F$.

Let t be an irrational number. Suppose that contrary to the claim, $F \succ E$. This can be assumed without loss of generality, because if $E \succ F$ we write $\pi(F) = t'\pi(E)$ for $t' = 1/t$.

We derive a contradiction. By Lemma 1, $\mathcal{C}^-(F)$ is not empty. By Lemma 3, there exists an event D such that $F \succ E \setminus D$, $D \subseteq \cup_{c \in \mathcal{C}^-(F)} E_c$, and $D \cap E_c$ is non-null for each $c \in \mathcal{C}^-(F)$. We denote $H_c = E_c \setminus D$. Let $\varepsilon = \min\{p_c(E_c \cap D) \mid c \in \mathcal{C}^-(F)\}$. Then, $\varepsilon > 0$ and we can choose a rational number k/n such that $t - \varepsilon < k/m < t$. Given this relation we have by the non-atomicity of the probabilities p_c an event $G \subseteq E$ such that $\pi(G) = (k/m)\pi(F)$. Moreover, for $c \in \mathcal{C}^-(F)$, we can choose G_c to satisfy $H_c \subseteq G_c$ where the difference is a non-null event.

As we have shown, $G \sim F$. Therefore, if $F \succ F_c$ then $G \sim F \succ F_c \sim G_c$. Thus, $\mathcal{C}^-(G) = \mathcal{C}^-(F)$, and similarly, $\mathcal{C}^+(G) = \mathcal{C}^+(F)$. We apply Lemma 2 to $X = E \setminus D$. The event X^- is $H = \cup_{c \in \mathcal{C}^-(G)} H_c \subset G^-$, and $X^+ = E^+ \supset G^+$. We conclude that $F \succ E \setminus D \succ G \sim F$, which is a contradiction. $\square$

5.4 From desirability to a relation in a Euclidian space

Using Proposition 3, we describe a binary relation on $\mathbb{R}^{\mathcal{C}}$. We use the notation $\succsim$ for both this relation and the relation on events, and call both desirability relations. No confusion will result.

Definition 7 Denote by $\mathbb{R}^{\mathcal{C}}_+$ the set of all points $x \in \mathbb{R}^{\mathcal{C}}$ such that $x \geq 0$ and $x \neq 0$. We define a relation on $\mathbb{R}^{\mathcal{C}}_+$ by $x \succsim y$ if there exist events E and F and positive numbers t and s such that $\pi(E) = tx$, $\pi(F) = sy$, and $E \succsim F$.

Note that if $x \succsim y$ then by Proposition 3, $E' \succsim F'$ for any pair of events E' and F' such that $\pi(E') = t'x$ and $\pi(F') = s'y$, for $t', s' > 0$.

Denote $\mathcal{M}(x) = \{y \mid y \succsim x\}$, $\mathcal{M}_+(x) = \{y \mid y \succ x\}$, $\mathcal{L}(x) = \{y \mid x \succsim y\}$, $\mathcal{L}_-(x) = \{y \mid x \succ y\}$, and $\mathcal{E}(x) = \{y \mid y \sim x\}$.

The next proposition addresses the convexity of these sets.

Proposition 4

1. *The relation $\succsim$ on $\mathbb{R}^{\mathcal{C}}_+$ is complete and transitive.*
2. *For each x, the sets $\mathcal{M}(x)$, $\mathcal{M}_+(x)$, $\mathcal{L}(x)$, $\mathcal{L}_-(x)$, and $\mathcal{E}(x)$ are convex cones.*

Proof

1. For x and y in $\mathbb{R}^{\mathcal{C}}$ there exist small enough positive t and s such that for some events E and F, $\pi(E) = tx$ and $\pi(F) = sy$. Since at least one of the relations $E \succsim F$ or $F \succsim E$ holds, it follows that at least one of $x \succsim y$ or $y \succsim x$ must hold. Suppose $x \succsim y$ and $y \succsim z$. Then there are events E, F, and positive numbers t_E and t_F, such that $\pi(E) = t_E x$, $\pi(F) = t_F y$, and $E \succsim F$. There are also events G and H, and positive numbers t_H and t_G, such that $\pi(G) = t_G y$, and $\pi(H) = t_H z$, where $G \succsim H$. Since $\pi(G) = t_G t_H^{-1} \pi(H)$, it follows by Propostion that $G \sim H$. Hence, $E \succsim H$ and therefore $x \succsim z$.
2. The sets in this part of the proposition are cones by the definition of $\succsim$. Consider the set $\mathcal{M}(x)$. To prove that it is convex it is enough to show that for

any $z, w \in \mathcal{M}(x)$, $z + w \in \mathcal{M}(x)$. Let G be an event such that $\pi(G) = rx$. For small enough $t > 0$ there are disjoint events E and F such that $\pi(E) = tz$ and $\pi(F) = tw$. Hence, $E \succsim G$ and $F \succsim G$. By Corollaries and 4, $E \cup F \succsim G$. But $\pi(E \cup F) = t(z + w)$ and thus $z + w \in \mathcal{M}(x)$. The proof for the rest of the sets is similar. □

Next, we discuss the topological properties of these sets. We denote by e_c the unit vector of the coordinate c, and write e_i for e_{c_i}.

Proposition 5 *For each $x \in \mathbb{R}_+^{\mathcal{C}}$:*

1. *the sets $\mathcal{M}_+(x)$ and $\mathcal{L}_-(x)$, are open subsets in $\mathbb{R}_+^{\mathcal{C}}$. If $x \neq e_1$ then $\mathcal{L}_-(x) \neq \emptyset$. If $x \neq e_n$ then $\mathcal{M}_+(x) \neq \emptyset$;*
2. *the sets $\mathcal{M}(x)$, $\mathcal{L}(x)$, and $\mathcal{E}(x)$ are closed subsets in $\mathbb{R}_+^{\mathcal{C}}$;*
3. *the interior of $\mathcal{E}(x)$ is empty.*

Proof

1. Let $y \in \mathcal{M}_+(x)$ and suppose that $\pi(E) = ty$ and $\pi(F) = sx$. We may assume without loss of generality that $ty_c < 1$ for each c. As $E \succ F$ we can apply axiom A3 of Small Events. Consider a consequence c. If $P_c(E_c) > 0$, then E_c is non-null, and we can find an element Π_i of the partition Π such that $\Pi_i \cap E_c$ is non-null. Let $D_c = E_c \cap \Pi_i$. Then $E \setminus D_c \succ F$. As $\pi(E \setminus D_c) = ty - p_c(D_c)e_c$, it follows that $y - t^{-1}p_c(D_c)e_c \succ x$. Thus, at a point y which is not on the face $y_c = 0$, we can decrease the c -coordinate and remain in $\mathcal{M}_+(x)$. Similarly, since $C \setminus E_c$ is non-null, per our assumption on ty, we can choose an element Π_i of the partition Π, such that $(C \setminus E_c) \cap \Pi_i$ is non-null. By setting $D_c = (C \setminus E_c) \cap \Pi_i$, we have $E \cup D_c \succ F$. In this way we show that $y + t^{-1}p_c(D_c)e_c \succ x$. Thus, we can increase the c-coordinate and remain in $\mathcal{M}_+(x)$. Since $\mathcal{M}_+(x)$ is convex, to prove that it is open it is enough to show that for each point y in $\mathcal{M}_+(x)$ an interval along the c-coordinate containing y is in $\mathcal{M}_+(x)$. If $x \neq e_n$, then $e_n \succ x$ and hence $\mathcal{M}_+(x)$ is not empty. The proof for the set $\mathcal{L}_-(x)$ is similar.
2. The sets $\mathcal{M}(x)$ and $\mathcal{L}(x)$ are the complements in $\mathbb{R}_+^{\mathcal{C}}$ of $\mathcal{L}_-(x)$ and $\mathcal{M}_+(x)$ correspondingly, and hence they are closed. The set $\mathcal{E}(x)$ is the intersection of $\mathcal{M}(x)$ and $\mathcal{L}(x)$ and hence closed.
3. Let $y \in \mathcal{E}(x)$. There exists c such that either $y \succ e_c$ or $e_c \succ y$. Suppose the first holds. We can assume without loss of generality that $y = \pi(E)$ and $y_c < 1$. Choose $F_c \subseteq C$, such that $F_c \cap E = \emptyset$ and $p_c(F_c) < \varepsilon$. Then $E \succ E \cup E_c$. This means that $y \succ y + \varepsilon e_c$, and therefore $y + \varepsilon e_c \notin \mathcal{E}(x)$. This shows that y is not in the interior of this set. The proof for the case $e_c \succ y$ is similar. □

For $x \notin \{e_1, e_n\}$, the three sets $\mathcal{M}_+(x)$, $\mathcal{L}_-(x)$ and $\mathcal{E}(x)$ form a partition of $\mathbb{R}_+^{\mathcal{C}}$. The first two are disjoint open convex cones. Since $\mathcal{E}(x)$ does not have an interior point, it is the closure of each of the first two sets. These two convex open sets can be separated by a hyperplane. Since 0 is in the closure of the separated sets, the hyperplane is an $(n-1)$-dimensional subspace $\mathcal{S}(x)$. As $\mathcal{E}(x)$ is the closure of both sets, it must be the intersection of $\mathcal{S}(x)$ with $\mathbb{R}^{\mathcal{C}}_{+}$. Since the two separated sets are open, $\mathcal{E}(x)$ contains an interior point of $\mathbb{R}^{\mathcal{C}}_{+}$. Thus we conclude:

Corollary 5 *For $x \notin \{e_1, e_n\}$, the set $\mathcal{E}(x)$ is the intersection of $\mathbb{R}^{\mathcal{C}}_{+}$ with an $(n-1)$-dimensional subspace, $\mathcal{S}(x)$. This intersection is of dimension $n-1$, that is, it contains interior points of $\mathbb{R}^{\mathcal{C}}_{+}$.*

5.5 Likelihood relation in the Euclidean space

Using the desirability relation of events we defined a likelihood relations $\succsim_H$ on events which are equally desirable. We now show how such relations are transformed to a relation in $\mathbb{R}^{\mathcal{C}}$.

For $v \not\sim x$ we define a relation $\succsim^*_v$ on $\mathcal{E}(x)$.

Definition 8 For $y, z \in \mathcal{E}(x)$, if $x \succ v$, then $y \succsim^*_v z$ when $y + v \succsim z + v$, and if $v \succ x$ then $y \succsim^*_v z$ when $z + v \succsim y + v$.

By axiom A5 of Persistency, if $u, v \not\sim x$ then $\succsim^*_u = \succsim^*_v$. We denote this relation which is independent of the choice of v, by $\succsim^*$. We study the following sets that are defined in terms of this relation.

For each $y \in \mathcal{E}(x)$, we define five subsets of $\mathcal{E}(x)$: $\mathcal{M}^*(y) = \{z \mid z \succsim^* y\}$, $\mathcal{M}^*_+(y) = \{y \mid z \succ^* y\}$, $\mathcal{L}^*(y) = \{z \mid y \succsim^* z\}$, $\mathcal{L}^*_-(y) = \{z \mid y \succ^* z\}$, and $\mathcal{E}^*(y) = \{z \mid z \sim^* y\}$.

First, we describe the convexity properties of these sets.

Proposition 6

1. *The relation $\succsim^*$ on $\mathcal{E}(x)$ is complete and transitive.*
2. *For each $y \in \mathcal{E}(x)$, the sets $\mathcal{M}^*(y)$, $\mathcal{M}^*_+(y)$, $\mathcal{L}^*(y)$, $\mathcal{L}^*_-(y)$, and $\mathcal{E}^*(y)$ are convex.*

Proof

1. Since either $y + v \succsim z + v$ or $z + v \succsim y + v$, it follows that either $y \succsim_v z$ or $z \succsim_v y$. Suppose $y \succsim_v z$ and $z \succsim_v w$. Then $y + v \succsim z + v \succsim w + v$ and therefore $y \succsim_v w$.
2. Let $z, w \in \mathcal{M}^*(y)$ and $\alpha \in (0, 1)$. Then for some v such that $x \succ v$, $z + v \succsim y + v$ and $w + v \succ y + v$. Therefore, $\alpha z + \alpha v \succsim y + v$, and $(1-\alpha)w + (1-\alpha)v \succsim y + v$. By intermediacy, $\alpha z + (1-\alpha)w + v \succsim y + v$. That is, $\alpha z + (1-\alpha)w \in \mathcal{M}^*(y)$. The proof for the rest of the sets is similar. □

The following lemma is used in the next proposition that describes the topological properties of these sets.

Lemma 4 *For all $y, z \in \mathcal{E}(x)$:*

1. *$z + y \succ^* y$;*

2. *if* $y \sim^* z$ *and* $t > 0$ *then* $ty \sim^* tz$.

Proof

1. Let $x \succ v$. By intermediacy, $z \sim y \succ y + v$. Therefore, $z + y + v \succ y + v$. Hence $z + y \succ^* y$.
2. If $y \sim^* z$ then for some v such that $x \succ v$, $y + v \sim z + v$. Therefore, $ty + tv \sim tz + tv$ and thus $ty \sim^* tz$. □

Proposition 7 *For each* $y \in \mathcal{E}(x)$:

1. *the sets* $\mathcal{M}^*_+(y)$ *and* $\mathcal{L}^*_-(y)$, *are non-empty open subsets in* $\mathcal{E}(x)$;
2. *the sets* $\mathcal{M}^*(y)$, $\mathcal{L}^*(y)$, *and* $\mathcal{E}^*(y)$ *are closed subsets in* $\mathcal{E}(x)$;
3. *the interior of* $\mathcal{E}^*(y)$ *in* $\mathcal{E}(x)$ *is empty*.

Proof

1. By Lemma 4, $y + \varepsilon y \succ^* y \succ^* y - \varepsilon y$ and thus $\mathcal{M}^*_+(y)$ and $\mathcal{L}^*_-(y)$ are not empty. This also shows that close enough to $\mathcal{E}^*(y)$ there are points not in this set, which proves 3. If $z + v \succ y + v$, then by Proposition 5 there is a ball B around $z + v$ such that for each $w \in B$, $w \succ y + v$. Therefore, there is a ball B' around y such that for each $w' \in B'$, $w' + v \succ y + v$. Thus $y \in B' \cap \mathcal{E}(x)$ which shows that $\mathcal{M}^*_+(y)$ is open. The proof for $\mathcal{L}^*_-(y)$ is similar.
2. The first two sets are complements of open sets, and the third is the intersection of the first two. □

5.6 Separation

By Propositions 6 and 7 we can separate $\mathcal{M}^*(y)$ and $\mathcal{L}^*(y)$ by a hyperplane. Since $\mathcal{E}^*(y)$ is the boundary of each of these sets it is contained in this hyperplane. As the separated sets are of dimension $n - 1$, $\mathcal{E}^*(y)$ is of dimension $n - 2$. Thus,

Corollary 6 *For* $y \in \mathcal{E}(x)$, *there exists a unique subspace* $L(x, y)$ *of dimension* $n - 2$ *such that* $\mathcal{E}^*(y) = (L(x, y) + y) \cap \mathcal{E}(x)$.

We next show in two steps that the space $L(x, y)$ is independent of x and y.

Proposition 8 *There exists an* $(n - 2)$*-dimensional subspace* L *such that for all* x *and* $y \in \mathcal{E}(x)$, $L(x, y) = L$.

We prove it with the next three lemmas. We first fix x and vary y.

Lemma 5 *For each* x *there exists* $L(x)$ *such that for all* $y \in \mathcal{E}(x)$, $L(x, y) = L(x)$.

Proof Let $y' \in \mathcal{E}(x)$. By the separation, the ray ty must intersect $\mathcal{E}^*(y')$, and thus, for some $t > 0$, $y' \sim^* ty$ and hence $\mathcal{E}^*(y') = \mathcal{E}^*(ty)$. By Lemma 4, $\mathcal{E}^*(ty) = t\mathcal{E}^*(y)$. But $t\mathcal{E}^*(y) = t[L(x, y) + y) \cap \mathcal{E}(x)] = (L(x, y) + ty) \cap \mathcal{E}(x)$. Thus, $L(x, y') = L(x, y)$. □

In order to show that $L(x)$ is independent of x we use the next lemma.

Lemma 6 *For each* x,y, *and* z, *if* $x \sim^* y$, *then* $x + z \sim^* y + z$.

Proof If $x \sim^* y$, then by definition $x \sim y$. Suppose $x \succ z$. As $x \sim^* y$ it follows that $x + z \sim y + z$. In order to show that $x + z \sim^* y + z$ it is enough to find some v such that $x + z \succ v$, and $x + z + v \sim y + z + v$. Indeed, take $v = z$, then by Intermediacy $x + z \succ z$, and as $x \sim^* y$, $x + (z + z) \sim y + (z + z)$. The proofs for the cases that $z \succ v$ and $z \sim z$ are similar. □

Lemma 7 *There exists* L *such that for all* x, $L(x) = L$.

Proof For x and x' choose $y \in \mathcal{E}(x)$ and $y' \in \mathcal{E}(x')$ such that $y' - y = z \in \mathbb{R}^{\mathcal{C}}_{+}$. By Lemma 6, $\mathcal{E}^*(y) + z \subseteq \mathcal{E}^*(y')$. But, $\mathcal{E}^*(y') = (L(x') + y') \cap \mathcal{E}(x')$, and $\mathcal{E}^*(y) + z$ is an $(n-2)$-dimensional subset of $L(x) + y + z = L(x) + y'$. Therefore, $L(x) = L(x')$. □

This completes the proof of Proposition 8.

Since L is of dimension $n-2$ there are many linear functionals p such that $pw = 0$ for all $w \in L$. By the definition of L, each such functional separates $\mathcal{M}^*(y)$ and $\mathcal{L}^*(y)$, and contains $\mathcal{E}^*(y)$ for every x and $y \in \mathcal{E}(x)$. The separating functional p is going to play the role of consequential probabilities. Therefore we need the following claim.

Proposition 9 *The functional* p *can be chosen to be a strictly positive probability vector.*

Proof Let p' be a separating functional. By Lemma 4, for fixed x and $y \in \mathcal{E}(x)$, and for any $w \in \mathcal{E}(x)$, $y + w \in \mathcal{M}^*_{+}(y)$. Therefore, $p'(y + w) \neq p'y$ and thus $p'w \neq 0$.

Since $\mathcal{E}(x)$ is the intersection of $\mathbb{R}^{\mathcal{C}}_{+}$ with a subspace $\mathcal{S}$ of dimension $n-1$, there exists a non-zero functional $q \in \mathbb{R}^{\mathcal{C}}$ such that for each $w \in \mathbb{R}^{\mathcal{C}}_{+}$, $qw = 0$ if and only if $w \in \mathcal{E}(x)$.

Consider the two-dimensional space $\alpha p' + \beta q$. We show that it contains a point in $\mathbb{R}^{\mathcal{C}}_{+}$. Suppose to the contrary that $\{\alpha p' + \beta q \mid \alpha, \beta \in \mathbb{R}\} \cap \mathbb{R}^{\mathcal{C}}_{+} = \emptyset$. Then, the two sets can be separated by a non-zero functional w. Since the first set is a subspace, $w(\alpha p' + \beta q) = 0$ for each α and β, and we can assume that $wr \geq 0$ for all $r \in \mathbb{R}^{\mathcal{C}}_{+}$ which implies that $w \in \mathbb{R}^{\mathcal{C}}_{+}$. By the separation, $wq = 0$ and $wp' = 0$. The first equality implies that $w \in \mathcal{E}(x)$. But then the second equation is impossible because we proved that $p'w \neq 0$ for each $w \in \mathcal{E}(x)$. Therefore, we can choose $p = \alpha p' + \beta q$ in $\mathbb{R}^{\mathcal{C}}_{+}$. By the definition of q, for every $z \in \mathcal{E}^*(y)$, $pz = \alpha p'z = \alpha p'y = py$, which shows that p vanishes on L.

To see that p is strictly positive, note that for e_c, $pe_c = p_c$. By Lemma 4, $e_c + e_c \succ^* e_c$ and therefore $2p_c > p_c$, which shows that $p_c > 0$. We can assume that p is normalized and therefore it is a strictly positive probability vector. □

5.7 The family of separating functionals

When $n = 2$ the dimension of L is 0. The probability vector p can be chosen in this case to be any vector $(a, 1-a)$ for $0<a<1$. We now assume that $n > 2$ and construct a basis for L.

Proposition 10 *For each $i = 2, \ldots, n-1$ there is a unique pair of positive numbers δ_i, η_i, such that the vector d^i, defined by $(d^i_{i-1}, d^i_i, d^i_{i+1}) = (\delta_i, -1, \eta_i)$ and $d_j = 0$ for all $j \notin \{i-1, i, i+1\}$, is in L. The vectors d^i form a basis of L.*

Proof For $i = 2, \ldots, n-1$, let $\mathbb{R}(i)$ be the subspace of $\mathbb{R}^{\mathcal{C}}$ spanned by e_{i-1}, e_i, and e_{i+1}, and $\mathbb{R}_+(i) = \mathbb{R}(i) \cap R^{\mathbb{C}}_+$. Since the dimension of L is $n-2$, the dimension of $L \cap \mathbb{R}(i)$ is at least 1, and it cannot be higher than 1 because then there would be $x, y \in \mathbb{R}(i)$ such that $x > y$ and $x - y \in L$, contrary to Lemma 4. Thus, $L \cap \mathbb{R}(i)$ is of dimension 1.

Choose two distinct points x and y in the interior of $\mathbb{R}_+(i)$ such $x - y \in L$. We show that $x_i \neq y_i$. Suppose to the contrary that $x_i = y_i$. Since $x - y \in L$, it follows that $p_{i-1}(y_{i-1} - x_{i-1}) + p_{i+1}(y_{i+1} - x_{i+1}) = 0$. Since $p > 0$, $y_{i-1} - x_{i-1}$ and $y_{i+1} - x_{i+1}$ are of different signs. But $e_{c_{i-1}} \succ x \succ e_{c_{i+1}}$ and thus by Lemma 2 either $y \succ x$ or $x \succ y$, which contradicts the assumption that $x \sim y$.

Thus, we can assume without loss of generality that $y_i < x_i$. Now, $p_{i-1}(y_{i-1} - x_{i-1}) + p_i(y_i - x_i) + p_{i+1}(y_{i+1} - x_{i+1}) = 0$, and since the middle term is negative, $p_{i-1}(y_{i-1} - x_{i-1}) + p_{i+1}(y_{i+1} - x_{i+1}) > 0$. Therefore it is impossible that $y_{i-1} - x_{i-1} \leq 0$ and $y_{i+1} - x_{i+1} \leq 0$. Also, as $e_{c_{i-1}} \succ y \succ e_{c_{i+1}}$, it is impossible that one difference is positive and the other is non-negative, because this would imply contrary to $x \sim y$, that either $y \succ x$ or $x \succ y$. Therefore both are positive. Let,

$$\delta_i = \frac{y_{i-1} - x_{i-1}}{x_i - y_i} \tag{2}$$

and

$$\eta_i = \frac{y_{i+1} - x_{i+1}}{x_i - y_i}. \tag{3}$$

Then $y - x = (x_i - y_i)d^i$. Since $x - y \in L$, it follows that $d^i \in L$. Since $L \cap \mathbb{R}(i)$ is a line, δ_i and η_i are uniquely determined.

Since the vectors $d^2, \ldots, d^{n-1}$ are $n-2$ independent vectors they are a basis of L. □

The following proposition is a corollary of the proof of Proposition 10.

Proposition 11 *The vector p is in L if and only if for each $i = 2, \ldots, n-1$, and x and y in $\mathbb{R}(i)$ that satisfy $x \sim y$ and $x \sim^* y$,*

$$\delta_i p_{i-1} + \eta_i p_{i+1} = p_i. \tag{4}$$

5.8 Utility

We now construct a utility vector $u = (u_c)$, where we write u_i for u_{c_i}. We say that u is *monotonic* if $u_i < u_{i+1}$ for $i = 1, \ldots, n-1$.

Proposition 12 *There exists a monotonic vector u such that the function*

$$\hat{u}(x) = \sum_c p_c x_c u_c / px$$

on $\mathbb{R}_+^{\mathcal{C}}$ *is constant on* $\mathcal{E}(x^0)$, *for each* $x^0 \in \mathbb{R}_+^{\mathcal{C}}$. *The vector u is uniquely determined up to transformations* $u \to \alpha(u_1 + \beta, u_2 + \beta, \ldots, u_n + \beta)$, *for* $\alpha > 0$.

Proof When $n = 2$, $\mathcal{E}(x^0)$ is simply the ray $\{tx^0 \mid t > 0\}$. Since $\hat{u}$ is homogeneous, the claim of the proposition holds for any monotonic vector (u_1, u_2). Assume now that $n > 2$.

Consider first $x \in \mathcal{E}^*(x^0)$. Since $px = px^0$, $\hat{u}(x) = \hat{u}(x^0)$ is equivalent to

$$\sum_c p_c (x_c - x_c^0) u_c = 0.$$

By Proposition 10, for small enough t, $x = x^0 + td^i \in \mathcal{E}^*(x^0)$. The last equality in this case is equivalent to:

$$\delta_i p_{i-1} u_{i-1} + \eta_i p_{i+1} u_{i+1} = p_i u_i. \tag{5}$$

Using Eqs. (4), (5) can be written as

$$\delta_i p_{i-1}(u_{i-1} - u_{i-1}) + \eta_i p_{i+1}(u_{i+1} - u_{i-1}) = p_i(u_i - u_{i-1}). \tag{6}$$

This gives rise to: $(u_{i+1} - u_{i-1})/(u_i - u_{i-1}) = p_i/(\eta_i p_{i+1})$. Denoting $\Delta u_i = u_i - u_{i-1}$ for $i = 2, \ldots, n$, Eq. (6) is $(\Delta u_{i+1} + \Delta u_i)/\Delta u_i = p_i/(\eta_i p_{i+1})$, or

$$\frac{\Delta u_{i+1}}{\Delta u_i} = \frac{p_i}{\eta_i p_{i+1}} - 1 = \frac{\delta_i p_{i-1}}{\eta_i p_{i+1}}, \tag{7}$$

where the right-hand side is positive. Thus, choosing arbitrarily $u_1 < u_2$, the rest of the coordinates of u are determined by 7, and as the Δu_i's are positive, u is monotonic. Obviously, a vector v solves (5) if and only if for some $\beta \in \mathbb{R}^{\mathcal{C}}$ and a positive α, $v = \alpha(u_1 + \beta, u_2 + \beta, \ldots, u_n + \beta)$.

Now, considering tx^0. Obviously, $\hat{u}(tx^0) = \hat{u}(x^0)$. Thus the function $\hat{u}$ is constant on $\cup_{t>0} \mathcal{E}^*(tx^0)$, which is $\mathcal{E}(x^0)$. □

Proposition 13 $x \succsim y$ *if and only if* $\hat{u}(x) \geq \hat{u}(y)$.

Proof In the previous proposition we constructed u such that if $x \sim y$ then $\hat{u}(x) = \hat{u}(y)$. It is enough now to show that $y \succ x$ if and only if $\hat{u}(y) > \hat{u}(x)$.

Denote by X^i the set of points in $\mathbb{R}_+^{\mathcal{C}}$ such that $x_k = 0$ for all $k \notin \{i, i+1\}$. Clearly, for $x \in X^i$, $\hat{u}(x) \in [u_i, u_{i+1}]$ and $e_{i+1} \succsim x \succsim e_i$. Let $X = \cup_{i=1}^{n-1} X^i$. We first prove the claim for points in X. Suppose $x, y \in X^i$. We can assume that $y_i = x_i$. By the definition of P_{i+1}, $y \succ x$ if and only if $y_{i+1} > x_{i+1}$. But this holds if and only if $\hat{u}(y) \geq \hat{u}(x)$.

Next, suppose that $y \in X^i$ and $x \in X^j$ for $j \neq i$. Then, $y \succ x$ if and only if $i+1 \leq j$ and it is not the case that $i+1 = j$ and $x \sim y \sim e_j$. But this is equivalent to $\hat{u}(y) > \hat{u}(x)$.

Observe now that for every $x \in \mathbb{R}_+^{\mathcal{C}}$ there exists a point $x' \in X$ such that $x' \sim x$. Indeed, there exists i such that $e_{i+1} \succsim x \succsim e_i$. Consider the sets $\mathcal{M}(x) \cap X^i$ and $\mathcal{L}(x) \cap X^i$. By Propositions 4 and 5 these are closed cones. The first contains e_i and the second e_{i+1}. Therefore there exist x' in X^i which belong to both. Thus $x' \sim x$. Now, $x \succ y$ if and only if $x' \sim y'$, which is equivalent to $\hat{u}(x') > \hat{u}(y')$. But, $\hat{u}(x') = \hat{u}(x)$ and $\hat{u}(y') = \hat{u}(y)$, which completes the proof. □

5.9 Proofs of Theorems 1–4

To complete the proof of Theorem 1* we define a probability P on Σ by $P(E) = p\pi(E) = \sum_c p_c P_c(E_c)$. Note, that as $p > 0$, an event E is P-null if and only if $\pi(E) = 0$, which holds, by Proposition 2, if and only if E is null. Now, $\sum_{c_i} u_i P(E \mid C_i) = \hat{u}(\pi(E))$. Since $E \succsim F$ if and only if $\pi(E) \succsim \pi(F)$, (P, u) represents $\succsim$ on Σ by Proposition 13.

Proof of Theorem 1 To prove the "only if" part of Theorem 1 we construct a new state space $(\hat{\Omega}, \hat{\Sigma})$, a new set of consequences $\mathbb{C}$, and a new relation $\hat{\succsim}$ on $\hat{\Sigma}$. The set $\hat{\Omega}$ is obtained by eliminating from Ω all events C_i that are null. The σ-algebra $\hat{\Sigma}$ consists of the events in Σ which are subsets of $\hat{\Omega}$. For $\mathcal{C}$, we partition the set of consequence for which C_i is non-null into equivalence classes such that c_i and c_j belong to the same class if $C_i \sim C_j$. The consequences in $\mathcal{C}$ are these equivalence classes.

We need to show that $\mathcal{C}$ has at least two points, that is that there are i and j such that C_i and C_j are non-null and $C_i \succ C_j$.

Let I be the set of indices i such that C_i is non-null. The set I is not empty, because otherwise, $\Omega = \cup_i C_i$ is null, and hence all events are null, contrary to Non-degeneracy. Suppose that all the events C_i with $i \in I$ are similar. Let E be a non-null event. For each $i \notin I$, E_{c_i} is null, and hence, $E \sim \cup_{i \in I} E_{c_i}$. For some indices $i \in I$, E_{c_i} must be non-null. Let I^* be the subset of I of such indices. Then, $E \sim \cup_{i \in I^*} E_{c_i}$. Choose $i^* \in I^*$. Then by Corollary, $E \sim E_{c_{i^*}}$. By axiom A6 of Consequence Events, $E \sim C_{i^*}$. Since this holds for all non-null events E, and all the C_{i^*} are similar, all non-null events are similar, contrary to Non-degeneracy.

Finally, the relation $\hat{\succsim}$ is the restriction of $\succsim$ to the events in $\hat{\Sigma}$. We skip the simple proof that $\hat{\succsim}$ satisfies axioms A1–A5 as well as Assumptions 1 and 2. By Theorem 1* there exists a pair $(\hat{P}, \hat{u})$ that represents $\hat{\succsim}$. We define a probability P on Σ by setting $P(E) = \hat{P}(E \cap \hat{\Omega})$. The utility u is defined arbitrarily on c_i that correspond to null C_i, and for all other c_i, $u(c_i) = \hat{u}(\hat{c}_j)$ where $\hat{c}_j$ is the equivalence class of c_i. We omit the straightforward proof that (P, u) represents $\succsim$. □

Proof of Theorem 2 Assume that $\succsim$ satisfies the said properties and (P, u) represents $\succsim$. We show that the conditional probability $P(\cdot \mid C)$ represents the qualitative probability relation $\gtrapprox$ in Definition 5. Since, by Proposition 1 there exists a unique probability on Σ_c that represents this relation, it follows that the conditional parts of probabilities in $\mathcal{P}(\succsim)$ are the same.

Consider an event $A \subseteq C$ and event H such that $H \cap C = \emptyset$. Then, the expected utility given $A \cup H$ is

$$\frac{P(C)P(A \mid C)u_c + \sum_{c' \neq c} P(C')P(H \mid C')u_{c'}}{P(C)P(A \mid C) + \sum_{c' \neq c} P(C')P(H \mid C')}. \tag{8}$$

Choose H such that $C \succ H$ (if there is none, we choose H such that $H \succ C$ and the argument is similar). Then u_c is greater than the expected utility given H. It follows that the derivative of (8) with respect to $P(A \mid C)$ is positive. Thus, For $A, B \subseteq C$, $A \cup H \succsim B \cup H$, which is equivalent to $A \gtrapprox B$, holds if and only if $P(A \mid C) \geq P(B \mid C)$.

A probability vector p is a consequential part of some $P \in \mathcal{P}(\succsim)$ if and only if it is a positive solution of the $n-2$ equations in (4). The set of positive solutions of these equations in the simplex is a maximal interval. Dividing Eq. (4) by p_i we obtain for $i = 2, \ldots, n$,

$$r_i = \frac{1 - \delta_i / r_{i-1}}{\eta_i}, \tag{9}$$

where $r = \rho(p)$. The function $(1 - \delta_i / x)/\eta_i$ is monotonic in $x > 0$. Thus, if q is in the said interval, $s = \rho(q)$, and $r_1 > s_1$, then $r_2 > s_2$, which implies that $r_3 > s_3$ and so on. That is, $p \gg q$. It is easy to check that the maximal interval that contains p and q is ordered.

Conversely, suppose that a family of probability $\mathcal{P}$ satisfies the two properties of the theorem. Let (P_i) be the unique conditional part of probabilities in $\mathcal{P}$. Let $p \neq q$ be two elements in the interval of consequential probabilities of $\mathcal{P}$, such that $q \gg p$. Consider the two equations $\lambda_i p_{i-1} + \eta_i p_{i+1} = p_i$ and $\lambda_i q_{i-1} + \eta_i q_{i+1} = q_i$ with variables δ_i and η_i. It is easy to see that these two equations have a unique solution and that it is positive. We define now a monotonic vector u by Eq. (7). The vectors p and u satisfy Eqs. (4) and (5). Let $P = \sum p_i P_i$ and let $\succsim$ be the desirability relation defined by the pair (P, u). Then, Eqs. (2) and (3) are satisfied and thus, the set of consequential probabilities of $\mathcal{P}(\succsim)$ is the set of positive solutions of Eq. (4). Since q is also in this set, $\mathcal{P} = \mathcal{P}(\succsim)$. $\square$

Proof of Theorems 3 and 4 Equation (7) shows that $\Delta u_{i+1}/\Delta u_i$ is uniquely determined by the consequential probability vector $(p_i) = (P(C_i))$, which means that u is determined up to a positive affine transformation. Moreover, it satisfies the equation in Theorem 4. $\square$

5.10 An example

We discuss here the desirability relation $\succsim$ from Example 1, which is represented by the pair (P, u) where P is the uniform probability distribution on the unit interval. We construct the family of all the probability-utility pairs that represent $\succsim$. We choose an example with three consequences because the case of two consequences is trivial. In this case all of $\Delta(\mathbb{C})$ is an interval ordered by optimism, and all utility functions are positive affine transformation of each other (with Assumption 1).

We project the P-non-null events in Σ to $\mathbb{R}^3_+$, the non-negative orthant of $\mathbb{R}^3$ without 0, by $\pi(E) = (P(E \mid C_i))_i$. By inequality (1) in Definition 4, the desirability relation between two events E and F depends only on $\pi(E)$ and $\pi(F)$. Moreover, if $\pi(E)$ and $\pi(F)$ are proportional then $E \sim F$. This last property makes it possible, just for convenience, to extend $\succsim$ to *all* of $\mathbb{R}^3_+$.

These claims on the relation $\succsim$ on $\mathbb{R}^3_+$ follow easily from the fact that the relation is defined by a probability–utility pair by inequality (1). In the proof of Theorem 1 we need to show that they follow from the axioms.

For $x \in \mathbb{R}^3_+$, let δ and η be the increase in x_1 and x_3 respectively, per a decrease of one unit of x_2, required for maintaining the same probability and the same conditional expected utility. Recalling that $P(C_i) = 1/3$ for $i = 1, 2, 3$, δ and η should satisfy:

$$(1/3)\delta + (1/3)\eta = (1/3)(1), \tag{10}$$

$$(1/3)(0)\delta + (1/3)(1)\eta = (1/3)(1)(1/2). \tag{11}$$

Equation (10) reflects the preservation of probability. Since, the probability is kept fixed, Eq. (11) reflects that preservation of the *conditional* expected utility. Observe also that these equations are the same for all x.

Equations (10) and (11) are derived from the given pair (P, u). In the proof of Theorem 1 we show how they can be derived from the axioms on $\succsim$.

The solution of (10) and (11) is $\delta = \eta = 1/2$. Thus, if the difference $x - y$ of two points x and y in $\mathbb{R}^3_+$ is in the direction $(1/2, -1, 1/2)$, the two points are similar, that is $x \sim y$ and have the same probability, that is $\sum_i (1/3)x_i = \sum_i (1/3)y_i$. In Fig. 1, the difference between $x = (1/4, 0, 1/2)$ and $y = (0, 1/2, 1/4)$ is in this direction. Therefore, the whole interval between x and y consists of points which are similar and have the same probability. By the homogeneity of similarity, the cone generated by x and y consists of similar points, and all the points in an interval parallel to the interval [x, y] in this cone have the same probability.

We now show the other pairs (Q, v) that represent the same relation $\succsim$. First, we know by Theorem 2 that $Q(\cdot \mid C_i) = P(\cdot \mid C_i)$ for each i. Thus, the projection of the Q-non-null events to $\mathbb{R}^3_+$ is the same as the projection of the P-non-null events. Also, since (P, u) and (Q, v) present the same desirability relation, the relation $\succsim$ on $\mathbb{R}^3_+$ is the same for both representations.

We show in the proof that having the same probability for two events that are similar is defined in terms of the desirability relation using axiom A5 of Persistency.

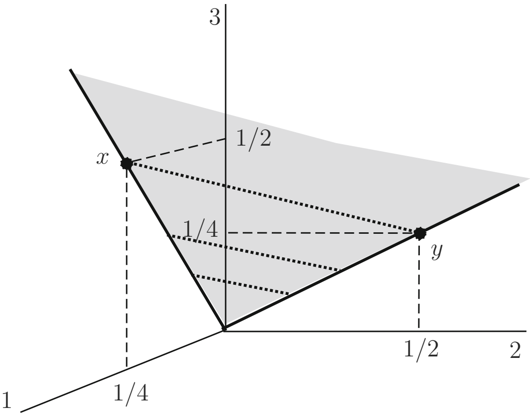

Fig. 1 Similarity and same probability

Since (Q, v) and (P, u) represent the same desirability relation, the picture of similarity and having the same probability for (Q, v) should look the same as the one in Fig. 1. Thus, the direction of having similarity and same probability should be $(1/2, -1, 1/2)$. Hence, the vector of consequential probability $q = (Q(C_i))_i$ and v should satisfy the following equations:

$$q_1(1/2) + q_3(1/2) = q_3(1), \tag{12}$$

$$q_1 v_1(1/2) + q_3 v_3(1/2) = q_2 v_2(1). \tag{13}$$

The positive probabilities that solve (12) form an open interval of probabilities between (2/3, 1/3, 0) and (0, 1/3, 2/3) as in Fig. 2. The point (1/3, 1/3, 1/3) with which we started is, of course, on this line. The closer the point in this interval is to (0, 1, 3, 2/3) the more optimistic it is. Thus, the likelihood ratio vector for (1/3, 1/3, 1/3) is (1, 1), while for (1/6, 1/3, 1/2) it is (2, 3/2) which dominates the first vector.

Fixing q that solves (12) and solving for v in (13) we find that $(v_3 - v_2)/(v_2 - v_1) = q_1/q_3 = (q_1/q_2)(q_2/q_3)$, which is the equality in Theorem 4.

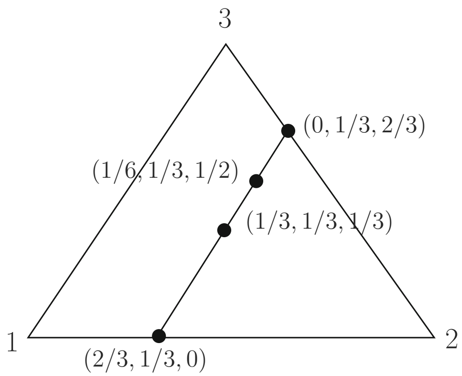

Fig. 2 The interval of consequential probabilities

Thus, v is uniquely determined by q, up to a positive affine transformation. Moreover, if q is more optimistic than p then the ratio of utility gains of v is dominated by that of u.

Acknowledgements Dov Samet acknowledges financial support of ISF Grant #722/18.

References

Ahn, D. S. (2008). Ambiguity without a state space. *The Review of Economic Studies, 75*, 3–28.

Aumann, R. (1987). Correlated equilibrium as an expression of Bayesian rationality. *Econometrica, 55*, 1–18.

Bolker, E. D. (1966). Functions resembling quotients of measure. *Transactions of the American Mathematical Society, 2*, 292–312.

Bolker, E. D. (1967). A simultaneous axiomatization of utility and subjective probability. *Philosophy of Science, 34*(4), 273–298.

Carlaw, K. I., Lipsey, R. G., & Webb, R. (2007). *Has the ICT Revolution Run its Course?, Discussion Papers dp07-18*. Department of Economics: Simon Fraser University.

Chai, J., Li, C., Wakker, P. P., Wang, T. V., & Yang, J. (2016). Reconciling savage's and Luce's modeling of uncertainty: The best of both worlds. *Journal of Mathematical Psychology, 75*, 10–18.

de Finetti, B. (1931). Sul significato soggettivo della probabilita. *Fundamenta Mathematicae, 17*, 298–329.

Gilboa, I. (1987). Expected utility with purely subjective non-additive probabilities. *Journal of Mathematical Economics, 16*, 65–88.

Jeffrey, R. (1965). *The logic of decision*. New York: McGraw Hill.

Jeffrey, R. (1983). *The logic of decision*. Chicago: The University of Chicago Press.

Luce, R. D., & Krantz, D. H. (1971). Conditional expected utility. *Econometrica, 39*, 253–271.

Machina, M. J., & Schmeidler, D. (1992). A more robust definition of subjective probability. *Econometrica, 60*, 745–780.

Savage, L. J. (1954). *The foundations of statistics*. New York: Wiley.

Stone, M. H. (1936). The theory of representation for Boolean algebras. *Transactions of the American Mathematical Society, 40*, 37–111.

Telpaz, A., & Levy, D. J. (2015). Consumers. Future choices. *Journal of Marketing Research, 52*, 511–529.

von Neumann, J., & Morgenstern, O. (1953). *Theory of games and economic behavior* (3rd ed.). Princeton: Princeton University Press.

Webb, R., Levy, I., Lazzaro, S. C., Rutledge, R. B., & Glimcher, P. W. (2019). Neural random utility: Relating cardinal neural observables to stochastic choice behaviour. *Journal of Neuroscience, Psychology, and Economics, 12*, 45–72.