Desirability relations in Savage's model of decision making*

Dov Samet and David Schmeidler

October 29, 2019

Abstract

Following Aumann (1987), we introduce the notion of *comprehensive* state of the world, which unlike Savage's state, describes a consequence. Comprehensive states result in Savage's model when an act is fixed. We study *desirability* which is a binary relation defined on events, that is, subsets of comprehensive states. One of our main results shows that a desirability relation that satisfies certain axioms can be represented by probability and utility, by taking conditional expected utility. Despite the difference of domains, some of our axioms are analogous to Savage's axioms. Some other axioms resemble axioms of Bolker on binary relation on events in a consequenceless model, that guarantee representation by ratio of two measures. Our axioms allow for endogenous derivation of consequences from a desirability relation, in contrast to Savage's model and similar models in which consequences, which are primitives of the theory, are assumed exogenously. The representation is not unique and we characterize all the pairs probability-utility that represent a desirability relation. We show that when desirability relations for a large set of acts satisfy some consistency axioms, then there exists a single pair that represents all of them.

Abstract

Following Aumann (1987), we introduce the notion of *comprehensive* state of the world, which unlike Savage's state, describes a consequence. Comprehensive states result in Savage's model when an act is fixed. We study *desirability* which is a binary relation defined on events, that is, subsets of comprehensive states. One of our main results shows that a desirability relation that satisfies certain axioms

*Acknowledgements will be added.

can be represented by probability and utility, by taking conditional expected utility. Despite the difference of domains, some of our axioms are analogous to Savage's axioms. Some other axioms resemble axioms of Bolker on binary relation on events in a consequenceless model, that guarantee representation by ratio of two measures. Our axioms allow for endogenous derivation of consequences from a desirability relation, in contrast to Savage's model and similar models in which consequences are primitives of the theory and assumed exogenously. The representation is not unique and we characterize all the pairs probability-utility that represent a desirability relation. We show that when desirability relations for a large set of acts satisfy some consistency axioms, then there exists a single pair that represents all of them.

1 Introduction

1.1 Interactive decision making

The decision maker (DM) in Savage's theory is facing various acts between which she has to chose. Each act may results in some possible consequences, but the DM is uncertain about which consequence will be realized. 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, an act can be identified with a function that assigns to each state a consequence. Obviously, a state of the world does not and should not describe the act of the DM about which she deliberates. In particular, the state does not specify any consequence.

The model proposed by Savage is for a single DM. But there is nothing in the model that excludes the possibility of other decision makers being there whose decisions may influence the consequence of the the DM act. In other words, the theory can be applied even when the DM is a player in a game. In this case, however, the description of the world should specify the acts of the other DM's in order to determine the consequence of the DM's chosen act.

An extension of the very well known example, taken from Savage (1954), illustrates this. Our DM 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 omelettere has three possible acts: breaking the egg into the bowl, breaking it into a saucer for inspection, or throw it away. The consequences describe the number of eggs in the omelette and the need to wash the saucer in case it was used.

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

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 omelettere will come again to buy eggs or not. These consequence depend on the acts of the omelettere. For example, if the omelettere throws away the egg, he'll 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 omelettere.

1.2 Comprehensive states

The two models that describe the decision problems of the omelettere and the egg seller are different models. In the model of each of the DMs the states specify the acts of the other DM, but not those of the DM whose decision is modeled. This lead Kadane and Larkey (1982) to conclude that subjective probability 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 agent in his own model we miss an important aspect of interaction, namely the reasoning of agents about each other choices. Interaction of reasoning requires one model for all agents. When interactive reasoning is studied in one model there are restrictions on the choices made by the agents that cannot be deduced if we analyze each one of them separately. This was demonstrated by Aumann (1987) who showed that common knowledge of rationality implies that the distribution of acts should be a correlated equilibrium. Such a conclusion cannot be derived by analyzing separately the choices of the interacting agents. Now, if all agents share the same model, namely the same set of states, then as we argued before, each state should describe the actions of each of the decision makers and hence also the consequences of these acts. We call a state that describes a consequence a *comprehensive* state.

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 implication of Bayesian rationality on the players' behavior, Aumann needs the players to have each 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 no consequences are associated with states. How can such probability and utility be derived in a comprehensive state space in Aumann (1987)?

This is the question we address in this paper. But here we are only 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 are studying a comprehensive state space of a single DM which does not require the introduction of knowledge structures. The results of this research will be used in subsequent papers to study the derivation of probability and utility in comprehensive state spaces of several players with knowledge structures.

1.3 Desirability

Probability and utility in a comprehensive state space are derived here from a relation over events in this space, which we call *desirability*. It is possible to give this relation few 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 DM happier, or more pleased, or more content than learning that the second happened. Alternatively, we can think of desirability as reflecting counterfactual choice. Although the DM cannot bring about one of two events, she can entertain the counterfactual situation in which she *can* chose the event that will obtain. Saying that one event is more desirable than another means that had she had the opportunity to chose, she would chose the first event to obtain. The following example demonstrates a comprehensive state space and a desirability relation on its events.

Consider Eve who contemplates the submission of her new paper to one of several equally reputed journals between which she is indifferent. A 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.

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 who handles 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 that handles the paper. Note that Alice handling the paper or Bob doing it, are not consequences. Eve's desire that the first event obtains 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 is accepted when Alice is the associate editor, Eve my find the latter event more desirable than the event the Bob is the associate editor.

There is a simple way to define a desirability relation on a comprehensive state space when a pair (P, u) of a probability P on the state space, and utility u of consequences. We define one event to be more desirable than another one, if the conditional expected utility on the first event is greater than the conditional expected utility on the second event. This desirability relation is said to be *represented* by the pair utility-probability that serves to define it. The purpose of this paper is to find conditions on the desirability relation, which are compatible with the intuitive meaning of this relation, that guarantee that it is represented by a pair probability-utility.

1.4 Choice and non-choice data

The purpose of economics is to explain and predict how individuals and other agents behave and choose in economic environments. Utilities and probabilities can together explain such behavior. In Savage's theory utilities and probabilities are derived from many observed choices. But there is no logical necessity that probabilities and utilities that *do explain choice and behavior* are derived from such observed behavior. An example is Aumann (1987) were acts do not vary, and probability, utility and common knowledge of rationality explain the choices made in a correlated equilibrium. In this model our observation is limited to specific choices made by the agents and not to many choices between pairs of actions. Yet, the explaining components can be derived, as we show here, from desirability which is not choice data. The purpose of economics is fully achieved, only that the data used to explain behavior comes from non-behavioral data.

But is desirability observable data? An important byproduct of the revolution in the technologies of communication and information in the last decades is the influx of reports by individuals about their lives which are accessible to other individuals and organizations. Part of these reports can be interpreted as statements of desirability in our sense. Firms including giants of technology, as well as state organizations, invest a lot of resources to collect, process, and use this information. Even the non behavioral data, about opinions, thoughts and desires is used to predict behavior, including economic behavior.[1] This paper is a step in laying theoretical foundations to the use of non-choice data to explain and predict economic behavior.

Other examples of non-choice data which can be used to predict economic choices are found in the literature on brain research and other somatic research. For example, Telpaz, Webb and Levy (2015) showed how the cheap test of electroencephalography (EEG) signals that result from just observing products without making any choice can predict the choices of these products at a later stage. Previous works showed how the expensive method of fMRI can achieve this goal. Such techniques can be also used, in principle, to replace reports in eliciting the desirability relation.

One of the advantages of collecting the non-choice data of desirability is the ability to use it in order to find out what the consequences that the decision maker is considering are. In a nutshell, an event is a consequence events if it is as desirable as its subevents and less or more desirable than some of its superevents. This is discussed in detail in subsection 2.2 when

[1]2 This usage seems to indicate that the problem of reliability of reports is not crucial. references.

axiom Com 5 of Consequence Events is presented. In contrast, within Savage's framework, or similar setups, it is impossible to infer from the choices made by the DM what consequence he has in mind. In order to translate the DM's observed choices into Savage's acts, one needs to *assume* a set of consequences without being able to confirm that these are indeed the DM's consequences.

1.5 Desirability on a fixed comprehensive state space

We sketch the gist of seven axioms, Com 1 to Com 7, on a desirability relation on a fixed **comprehensive** state space. These axioms appear to hold for the intuitive meanings of desirability discussed above. Not surprisingly, some of the axioms bear resemblance to Savage's axioms despite the different domains.

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

Axioms Com 1 - Com 3 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 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 (Com 1) requires that the relation is non-trivial. This axiom implies that there are non-null events, which makes the next axiom of *weak order* (Com 2) non-vacuous. The latter says that the desirability relation is a complete and transitive order on the non-null events. This type of axioms predate Savage and von Neumann and Morgenstern. One of the innovative axioms of Savage that play a crucial role in proving the existence of a probability is P6', the axiom of *Non-atomicity.* We give the same name to our (Comp 3) as it is essentially the same axiom. It says that the space can be partitioned into "small" events.[2]

Axiom Com 4 - Com 7 are special to desirability relations. Axiom Com 4 is a first glimpse into the notion of likelihood which is part of our intuition about desirability of events. It amounts to saying that if likelihood of consequences in two events are the same than the events are equally desirable. This is done, of course, in terms of the desirability relation. We illustrate

[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 preference relation on acts.

axiom Com 5 of *Consequence Events* in our example. Consider the following two events. The first is the event that the paper is accepted. The second is the conjunction of the first event with another one. Specifically the second event is the event that the paper is accepted *and* that Alice handled it. The second event is a subevent of the first and it is more informative. The axiom requires that these two events are equally desirable. The reason is simple. In both events the paper is accepted. The fact that in the second event it was handled by Alice, which is not a consequence, does not matter as long as the paper is accepted. The axiom requires in general that a consequence event is equally desired as any of its non-null subevents.

Note that if a desirability relation satisfies axiom Com 5 of Consequence Events then we can find out what the consequence events are and they do not have to be given exogenously. More concretely, a consequence event is one which is as desirable to the DM as all its subevents, but is less or more desirable than one of its superevents. That is, an event is a consequence event if more informative events are equally desirable but some less informative events are more or less desirable. In our example, we can conclude that acceptance is a consequence, by verifying that the for the DM acceptance is as desirable as any of its subevents but is more desirable, say, than the event that the paper is either acceptable or rejected.[3]

The axiom of *intermediacy* (Com 6) says that mixed news, good and bad, lie, in terms of its desirability, between the good and the bad news. Thus, in our example the event that the paper is either accepted or rejected is more desirable than the event of acceptance, but less desirable than the event of rejection.

Before we discuss the *persistency* (Com 7) we demonstrate how a likelihood relation between certain events can be deduced from the desirability relation. Let E and F be events that the DM equally desires, and H be an event disjoint of E and F and more desirable than both. On a first glance the event $E \cup H$ cannot be more desirable than $F \cup H$. But on a 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 will 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 definition

[3] In Savage's framework, consequence are given exogenously, and cannot be derived form the DM's preferences over acts.

of being more likely is independent of the event H that is used to define it.

Our first result is:

> *A desirability relation satisfies axioms Com 1 - Com 7 if and only if it is represented by of probability-utility pair* (P, u)*.*

A given desirability relation can be represented by more than one probability-utility pair. We characterize the structure of all representing pairs. First, all the probabilities are uniquely determined on the consequence event, That is,

- All the probabilities that are part of a representing pair have the same conditional probability on the consequence events.

Thus, these probabilities can differ only by assigning different probabilities to the consequence events themselves. We call such probability vectors *consequence probabilities*. We show that,

- The set of consequence probabilities is an interval in the finite vector space whose dimension is the size of the finite set of consequences.

We can say more about this interval. 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. Now,

- The consequence probabilities are arranged in the interval by optimism.

Finally,

- For any representing pair (P, u), the utility u is uniquely determined by P up to positive affine transformation, and the probability P is uniquely determined by u.

While the consequence probabilities are arranged linearly by optimism, the corresponding utilities are ordered by *contentment*. The utility u is more *content* than v if the gains from moving to a more desirable consequences, measured by the ratio of utility difference is higher in u than in v. If (P, u) and (Q, v) are representing pairs, and the consequence probability p is more optimistic than q, then the utility u is less content than v. Thus the optimism in consequence probabilities is set down by the contentment of the utility function.

1.6 Desirability when acts vary

So far we studied a desirability relation on one comprehensive state space determined by a fixed act. We now consider a family of acts and the corresponding comprehensive state spaces defined by them. We assume that on each of these spaces a desirability relation on events is defined. We impose axioms Act 1–Act 3 on the family of these binary relations. Axiom Act 1 of *common null event* requires that the set of null events is the same for each one of the binary relations. Axiom Act 2 of *common desirability* states that if two acts coincide on each one of two events, then the desirability relation between these events is the same in the two comprehensive state spaces defined by these two acts. Finally, axiom Act 3 of *Common Likelihood* asks that the likelihood relation defined in terms of the desirability relation is the same in all the comprehensive state spaces. We then show:

> *If a family comprehensive state spaces, equipped each with a desirability relation on events, satisfies the axioms Act 1-Act 3, then there exists a unique probability-utility pair that represents all the desirability relations.*

1.7 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 much the same they are defined on fields or sigma fields which are 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 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 and acts are the main features.

Based on a previous work, Bolker (1966), Bolker (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 Com 6 and Com 7. However, since their theory does not have consequences and acts it is impossible to vary the acts and require axioms Act 1–Act 3.[4]

In addition to the essential difference of having consequences and acts in our model, it 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 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, like 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), derive an 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. This enables us to demonstrate three features of the non-uniqueness in our model: (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.

A binary relation 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 numeric probability. By adding an axiom of non-atomicity Savage (1954) showed that a qualitative probability has a unique representation by an additive probability. Our axiom of non-atomicity, Com 3, is similar to Savage's, and we use his result to prove the existence of a probability on each consequence event.

Kreps (1979) studied preference over subsets of menus. Gul, Lipman and Pesendorfer (2001) and Dekel, Lipman and Rustichini (2001) followed suit with preferences over subset of lotteries. Their axioms are different than

[4] Jeffrey (1983) is a new version of Jeffrey (1965) in which Bolker's contribution is presented and discussed. See also Jeffrey (2004) for a summary of the history of their notion of desirability.

ours and their presentation theorems differ in form and contend from ours. Ahn (2008) studied also a preference over lotteries, and 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 some integral on utility divided by a probability. A representation by ratio similar in form to our result, in a risk context, is found in Chew (1983), where weighted utility has the form of a ratio of utility functions which are linear in 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 not a function on the whole state space but only on an event in this space.

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 $f\colon \Omega \to \mathcal{C}$ that specifies a consequence in each state.

For a fixed act f we refer to (Ω, Σ, f) as a *comprehensive state space* which 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 by f.

Fixing a 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 consider below Axioms Com 1–Com 7 ('Com' for comprehensive) 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, the symmetric difference of the two events.[5]

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

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 $(E'\Delta E) \cup (F'\Delta F) \subseteq N$.*

An immediate corollary of the 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 some 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 as E and F.*

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 observe 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 also $N \cup M \in \Sigma^0$;*

3. *If $N \in \Sigma^0$ and event M satisfies $M \subseteq N$, then $M \in \Sigma^0$.*

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

For part 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. $\square$

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 is possible that all events are null. However, we next show that if this relation is not trivial, then there must be some non-null events.

Claim 2. *If there are vents 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, $E \sim F$, contrary to our assumption that $E \succ F$. $\square$

Finally, it is easy to show that,

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 assumption is assumed also in Savage (1954) as well as in the axioms of qualitative probability in de Finetti (1931).

Com 1. (Non-degeneracy) *There are events E and F such that $E \succ F$.*

Axiom Com 1 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 $\succsim$ is defined only on pairs of non-null event, and on these events it is a weak order, namely complete and transitive. Since we are going to compare the desirability of only non-null events, we could allow, by Claim 3, that $\succsim$ can be defined on pairs that either both are, or one of them is null and require only that the restriction of $\succsim$ to the non-null event is a weak order.

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

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 effect the given relation. This axiom is a slight variation of property P6' in Savage (1954).

Com 3. (Non-atomicity) *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 introduce some notations. 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.

In the penultimate paragraph of subsection 1.3 we explained that the desirability of an event E depends on the likelihood of the subevents E_c. Thus, if the likelihood of the events E_c in E and the events F_c in F are the same, then E and F should be equally desirable. But to verify this equality of likelihood it is enough to check them for pairs. More specifically, for any pair of consequences c and d, it is enough to show that either $E_c \cup E_d$ and $F_c \cup F_d$ are both null, or else both are non-null and the likelihood of E_c and E_d in $E_c \cup E_d$ is the same as the likelihood of F_c and F_d in $F_c \cup F_d$. The latter condition is equivalent to saying that the two unions are equally desirable. Thus we can capture some aspects of likelihood purely in terms of desirability.

Com 4. (Pairs) *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 $E \sim F$.*

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 then any additional information is irrelevant to desirability. Hence the following axiom:

Com 5. (Consequence Events) *For any consequence c and a non-null event $E \subseteq C$, $E \sim C$.*

The following axiom formalizes the idea that a mixture of good news and 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.

Com 6. (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.*

As we have said, desirability of events is determined by the way they are related to the likelihood of consequences. Consider in the previous example the events $A =$ *acceptance* and $B =$ *acceptance and Alice handles the paper*, which by axiom Com 5 of Consequence Events are equally desirable. 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, then we expect that the relation of desirability between $A \cup G$ and $B \cup G$ would 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 such an implication to formally define a restricted concept of likelihood relation.

Definition 2. *Suppose that $A \sim B$, and 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 content of the next axiom which is in the spirit of the impartiality property in Bolker (1967).

Com 7. (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 also A is at least as likely as B according to H.*

We can now define for two equally desirable events A and B, the relation A is *at least as likely as* B. This relation holds if A is at least as likely as B according to *some* G. By axiom Com 7 of Persistency, this likelihood relation is well defined.

2.3 Consistency of desirability relations

We will prove that a desirability relation for a given act f is represented by a probability-utility pair. However this representation is not unique. We consider next desirability relations for a family of acts and formulate four axioms concerning this family that guarantee the existence of a unique probability-utility pair that represents all the desirability relations in this family.

We say that an act f is *full*, if for each consequence c, $C^f = f^{-1}(c)$ is non-null. Let $\mathcal{F}$ be a set of acts that contains all full acts. For each $f \in \mathcal{F}$ let $\succsim^f$ be a desirability relation on (Ω, Σ, f). We denote $\mathcal{D} = \{\succsim^f \mid f \in \mathcal{F}\}$. From now on we tag desirability relations in $\mathcal{D}$, as well as consequence events, and representing probabilities and utilities with a superscript of the act for which they are defined.

Act 1. (Common Null Events) *All the desirability relations in $\mathcal{D}$ have the same set of null events.*

By axiom Act 1 of Common Null Events we can refer to null or non-null events without specifying a desirability relation in $\mathcal{D}$.

The following axiom requires that the desirability relation between two events is independent of the value of the acts outside these events. For two acts f and g and an event E in Ω, we write $f =_E g$ if $f(\omega) = g(\omega)$ for each $\omega \in E$.

Act 2. (Common Desirability) *Let A and B be non-null such that $f =_A g$ and $f =_B g$. Then, $A \succsim^f B$ if and only if $A \succsim^g B$.*

Using axiom Com 7 of Persistency, we defined at the end of subsection 2.2, the relation of being at least as likely for two equally desirable events. This relation is defined for a given desirability relation. The next axiom requires that the relation of being at least as likely is the same for all $\succsim^f$ in $\mathcal{D}$. Since the relation is defined only on equally desirable events, the requirement that the relation is the same for $\succsim^f$ and $\succsim^g$ in $\mathcal{D}$ can be applied only to two events that are equally desirable for both $\succsim^f$ and $\succsim^g$.

Act 3. (Common Likelihood) *Let A and B be non-null events such that $A \sim^f B$ and $A \sim^g B$. If A is as at least as likely as B for $\succsim^f$, then A is as at least as likely as B for $\succsim^g$.*

3 The main theorems

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

Definition 3. (Representation) *Consider a pair (P, u), where P is a non-atomic probability on (Ω, Σ, f) and $u \colon \mathcal{C} \to \mathbb{R}$. 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. That is $u \mapsto \alpha u + \beta$ where $\alpha > 0$.

Theorem 1. *For a comprehensive state space* (Ω, Σ, f)*, a relation* $\succsim$ *on* Σ *satisfies axioms* Com 1–Com 7 *if and only if there exists a pair* (P, u) *that represents it.*

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 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, $u \colon \mathcal{C} \to \mathbb{R}$, 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 $i = 1 \ldots 3$, 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 it represents. The null events of $\succsim$ are the P-null-events. Note, that the conditional expectation given E is determined by the x's. Thus, in particular, if two events have the same conditional probability given each C_i, then they are similar.

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.

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 this inequalities describe optimism follows from Assumption 2. If $p \gg q$, then for each two consequences the likelihood of the preferred one is higher in p than in q. Let $\rho(p)$ be the $n-1$ dimensional vector defined by $\rho_i(p) = p_{i+1}/p_i$ for $i = 1, \ldots n-1$. We say that p *likelihood-ratio dominates*[6] 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 each $\alpha > \alpha'$, $\alpha p + (1-\alpha)q \gg \alpha' p + (1-\alpha')q$. The interval is *maximal* if p and q are on the boundary of the simplex.

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* Com 1–Com 7 *and Assumptions* 1,2 *if and only if:*

1. *The conditional parts of the probabilities in $\mathcal{P}$ are the same. 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.*

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

Finally, we characterize the utilities in the representing pairs.

Theorem 3. *For every $P \in \mathcal{P}(\succsim)$, there exists a unique utility u, 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. Denote $u_i = u(c_i)$ 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.

Roughly speaking, being more optimistic means giving higher probability to more desirable consequences, and being more content means giving less utility to 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 (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 each C_i is P_i. Thus, $Q = q_1 P_1 + q_2 P_2 + q_3 P_3$ 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 verified 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 Theorem 4.

In Section 5, we show how to compute the maximal interval of probability vectors that are ordered by optimism guaranteed by Theorem 2.

So far we have dealt with the representation of desirability for a given comprehensive state space with a given fixed act. We now vary the act in some family of acts $\mathcal{F}$ that contains all full acts. Each act in the family defines a comprehensive state space. For each $f \in \mathcal{F}$ the agent has a desirability relation $\succsim^f$ on the associated comprehensive state space. We denote the family of these desirability relation be $\mathcal{D}$. When the families $\mathcal{F}$ and $\mathcal{D}$

satisfy axioms Act 1–Act 3, then there exists a single probability-utility pair that represents all the desirability relations in this family.

Theorem 5. *The families of acts $\mathcal{F}$ and binary relations $\mathcal{D}$, satisfy axioms* Act 1–Act 3 *if and only if there exists a unique pair (P, u) that represents $\succsim^f$ for all acts $f \in \mathcal{F}$, where u is uniquely determined up to a positive affine transformation.*

Example 3. Consider the state space of Example 1, and the pair (P, u) in this example. This pair defines a preference relation over all acts by taking the expected utility for each act. Moreover, by Savage (1954) this pair is the unique pair that represents this preference relation. Theorem 5 shows that this pair can be uniquely determined by desirability relations rather than a preference relation on acts.

Let $\mathcal{F}$ be the set of all acts f such that for each i, $f^{-1}(c_i)$ is P-non-null. For each $f \in \mathcal{F}$, the pair (P, u) defines a desirability relation $\succsim^f$ on events in the comprehensive state space (Ω, Σ, f). The family of these relations satisfies axioms Act 1–Act 3. By Theorem 5, the pair (P, u) is the only pair that satisfies these axioms.

4 Proofs

4.1 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 the 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* Com 1–Com 7 *and Assumptions* 1 *and* 2 *if and only if there exists a pair (P, u) that represents it, such that for $i = 1, \ldots, n$, $P(C_i) > 0$ and $u(c_n) > u(c_{n-1}) > \cdots > u(c_1)$.*

In subsection 4.2 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 define a likelihood relation on a family of similar events. Since, by axiom Com 5 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 Com 3 of Non-atomicity it follows by a theorem of Savage that there exists a unique non-atomic probability on Σ_c, which represents the qualitative likelihood relation on Σ_c.

In subsection 4.3 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 subsection 4.4, the desirability relation on events to a relation on the positive orthant of $\mathbb{R}^{\mathcal{C}}$. We show that the sets defined by this latter relation are convex, and characterize their topological properties.

In subsection 4.5 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 subsection 4.6 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 subsection 4.7 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 subsection 4.2, the probabilities derived in subsection 4.5, and the utility derived in subsection 4.8, we go back to the desirability relation and prove Theorems 1-4. In the last subsection we prove Theorem 5.

4.2 The conditional probability over consequences

The following are three immediate corollaries of axioms Com 6 of Intermediacy and Com 2 of Weak Order. The first is not only a corollary of the two axioms, but combined with axiom Com 2 implies axiom Com 6.

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 $E^1 \sim E^2 \sim \cdots \sim E^k$, 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 $G \cap C = \varnothing$ 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 4. *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 Com 5 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 $B \cap (A \cup A') = \varnothing$.

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

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, $A_1 \cup G \succsim A_2 \cup G \succsim$ 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 Com 6 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 2 that $C \sim B \succ B \cup G$. Also $(B \cup G) \cap (A \cup A') = \varnothing$. Thus, by axiom Com 7 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, $A \cup G \succsim A' \cup G$ iff and only if $A \cup B \cup G \succsim A' \cup B \cup G$.

If A is non-null, then by Corollary 2, $A \cup G \succ G = \varnothing \cup G$. Hence it is not the case that $\varnothing \cup G \succsim A \cup G$, and therefore $A > \varnothing$. In particular, $C > \varnothing$. If A is null then $A \cup G \sim \varnothing \cup G$. Which show that for all A, $A \gtrapprox \varnothing$. 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 Com 3 of Non-atomicity 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 5. *Let $\pi \colon \Sigma \to \mathbb{R}^{\mathcal{C}}$ 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 $N_c \approx \varnothing$ 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 3 $E = E_c \cup (E \setminus C) \sim C$ and similarly $E' \sim C$ and we are done. Otherwise, $E \setminus C \not\sim C$. Now, $E_c = (E_c \cap E'_c) \cup N^1_c$ for some $N^1_c \subseteq N_c$. Since $P_c(N^1_c) = 0$, it follows by axiom Com 7 of Persistency, that $E = (E_c \cap E'_c) \cup N^1_c \cup (E \setminus C) \sim (E_c \cap E'_c) \cup (E \setminus C)$. 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, $F \sim F'$. □

4.3 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 $\mathcal{C}^+(G)$ 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, $G = \cup_{c \in C^+(G)} G_c \succ G$. □

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 $X \succ G$. □

Next, we describe a simple result of axiom Com 3 of Non-atomicity. 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 Com 3, 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 $E^1, ..., E^m$ 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 $E^1, ..., E^k$. Since $F \succ E \setminus D^k$, we can apply axiom Com 3 of Non-atomicity 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 Com 7 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 Com 4 of Pairs 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. $\square$

4.4 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 6. *Denote by $\mathbb{R}^{\mathcal{C}}_+$ the set of all point $x \in \mathbb{R}^{\mathcal{C}}$ 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 3 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

3 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. $\square$

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 $\mathcal{L}_-(x) \neq \varnothing$. If $x \neq e_n$ then $\mathcal{M}_+(x) \neq \varnothing$;*
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 Com 3 of Non-atomicity. 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. Denote $D_c = E_c \cap \Pi_i$. Then $E \backslash D_c \succ F$. As $\pi(E \backslash D_c) = ty - p_c(D_c)e_c$, 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 $F_c \cap E = \varnothing$ 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. $\square$

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 $\mathbb{R}^{\mathcal{C}}_{+}$.*

4.5 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 7. *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 Com 7 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. $\square$

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 $ty \sim^* tz$. $\square$

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. $\square$

4.6 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. □

4.7 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 are $x, y \in \mathbb{R}(i)$ 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, $p_{i-1}(y_{i-1} - x_{i-1}) + p_{i+1}(y_{i+1} - x_{i+1}) > 0$. Thus 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 $d^2, \dots, d^{n-1}$ are $n-2$ independent vectors they are a basis of L. $\square$

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 $i = 2, \dots, n-1$, and x and y in $\mathbb{R}(i)$ that satisfy $x \sim y$ and $x \sim^* y$,*

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

4.8 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 $i = 1, \dots, n-1$.

Proposition 12. *There exists a monotonic vector u such 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 $u \to \alpha(u_1 + \beta, u_2 + \beta, \dots, u_n + \beta)$, 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 equation (4), equation (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 $i = 2, \ldots, n$, Equation (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 point 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. □

4.9 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 $\hat{\mathbb{C}}$, 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 $\hat{\mathcal{C}}$, 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 $\hat{\mathcal{C}}$ are these equivalence classes.

We need to show that $\hat{\mathcal{C}}$ 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 3, $E \sim E_{c_{i^*}}$. By axiom Com 5 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 Com 1–Com 7 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 $\succsim$. $\square$

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 4. 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 form a maximal interval. Dividing equation (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 equation (7). The vectors p and u satisfy equations (4) and (5). Let $P = \sum p_i P_i$ and let $\succsim$ be the desirability relation defined by the pair (P, u). Then, equations (2) and (3) are satisfied and thus, the set of consequential probabilities of $\mathcal{P}(\succsim)$ is the set of positive solutions of equation (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. $\square$

4.10 Proof of Theorem 5

Proposition 14. *There exists a unique probability measure P on (Ω, Σ) such that for each $f \in \mathcal{F}$, C_i^f which is non-null, and $P^f \in \mathcal{P}(\succsim^f)$,*

$$P^f(\cdot \mid C_i^f) = P(\cdot \mid C_i^f).$$

Proof. Let $\mathcal{E}^+$ be the set of non-null events with non-null complements. Let $H \in \mathcal{E}^+$, and f and g be acts such that $H = C_i^f$ and $H = C_j^g$. Let $A, B \subseteq H$. By axiom Com 5 of Consequence Events, $A \sim^f B$ and $A \sim^g B$. Thus, by axiom Act 3 of Common Likelihood, $\gtrapprox_i^f = \gtrapprox_j^g$ which implies that $P_{c_i}^f = P_{c_j}^g$. We denote this probability on H, which is independent on the consequence, and the act by P_H.

Let $H \subseteq G$ be events in $\mathcal{E}^+$. We show that $P_H(\cdot) = P_G(\cdot \mid H)$. Let f be an act such that $H = C_i^f$ and g an act such that $G = C_i^g$. For $A, B \subseteq H$, again apply axiom Com 5 of Consequence Events and axiom Act 3 of Common Likelihood to conclude that $A \gtrapprox_i^f B$ if and only in $A \gtrapprox_i^g B$, which means that $P_G(A) \geq P_G(B)$ if and only if $P_H(A) \geq P_H(B)$. But this means that $P_H(\cdot) = P_G(\cdot \mid H)$.

We complete the proof by showing that there exists a unique non-atomic probability P on Σ such that for each $H \in \mathcal{E}^+$, $P_H(\cdot) = P(\cdot \mid H)$.

Let (A, B, X) be a partition of Ω into three non-null events. Then the three events, and the union of each two of them, are all in $\mathcal{E}^+$. We show that $P_{A\cup X}$ and $P_{B\cup X}$ determine P_H for each $H \in \mathcal{E}^+$. Obviously, $P_{H\cap A}(\cdot) = P_{A\cup X}(\cdot \mid H \cap A)$, $P_{H\cap B}(\cdot) = P_{B\cup X}(\cdot \mid H \cap B)$, and $P_{H\cap X}(\cdot) = P_{A\cup X}(\cdot \mid H \cap X) = P_{B\cup X}(\cdot \mid H \cap X)$. It remains to show that $P_{A\cup X}$ and $P_{B\cup X}$ determine $P_H(H \cap A) = \alpha$, $P_H(H \cap B) = \beta$, and $P_H(H \cap X) = 1 - (\alpha + \beta)$. Let $P_{A\cup X}(H \cap A) = p$ and $P_{B\cup X}(H \cap B) = q$. Assume first that $H \cap X$ is non-null. In this case, $p < 1$ and $q < 1$. Then $\alpha/(1 - (\alpha + \beta)) = p/(1 - p)$ and $\beta/(1 - (\alpha + \beta)) = q/(1 - q)$. These two equations determine $\alpha = (p - pq)/(1 - pq)$ and $\beta = (q - pq)/(1 - pq)$. If $H \cap X$ is null, then let $X' \subset X$ be a non-null event such that $X \setminus X'$ is also non-null. Thus, $H \cup X' \in \mathcal{E}^+$. Now, $P_{H\cup X'}$ is determined, and as $H \subseteq H \cup X'$, P_H is determined.

It is enough now to show that there exists a probability P such that $P_{A\cup X}(\cdot) = P(\cdot \mid A \cup X)$, and $P_{B\cup X}(\cdot) = P(\cdot \mid B \cup X)$. Denote $p = P_{A\cup X}(A)$

and $q = P_{B\cup X}(B)$. Let $P = \alpha P_A + \beta P_B + (1-(\alpha+\beta))P_X$. Then, for the conditionals of P on $A\cup X$ and $B\cup X$ to be as desired, α and β should satisfy the same equations as above, which have indeed a unique solution. □

Proposition 15. *The probability P is the unique element in $\cap_{f\in\mathcal{F}}\mathcal{P}(\succsim^f)$.*

Proof. To prove that P is in $\cap_{f\in\mathcal{F}}\mathcal{P}(\succsim^f)$, we fix $f \in \mathcal{F}$ and omit all superscripts referring to this act. We use the machinery of subsection 4.7, and make the assumption made there that f is a full act and $C_n \succ C_{n-1}, \succ \cdots \succ C_1$. The argument for the general case is similar to the one used in the proof of Theorem 1.

By Proposition 14, the projection of Σ to $\mathbb{R}^{\mathcal{C}}$ is given by $\pi(E) = (P(E \mid C_i)) = (P(E\cap C_i)/P(C_i))$. We need to show that $p = (P(C_i))$ satisfies equation (4) in Proposition 11. Let x and y be in $\mathbb{R}(i)$ such that $x = \pi(E)$, $y = \pi(F)$, where $x \sim y$ and $x \sim^* y$. By equations (2) and (3),

$$\delta_i = \frac{y_{i-1}-x_{i-1}}{x_i - y_i} = \frac{P(F\cap C_{i-1}) - P(E\cap C_{i-1})}{P(E\cap C_i) - P(F\cap C_i)}\frac{P(C_i)}{P(C_{i-1})}, \tag{10}$$

$$\eta_i = \frac{y_{i+1}-x_{i+1}}{x_i - y_i} = \frac{P(F\cap C_{i+1}) - P(E\cap C_{i+1})}{P(E\cap C_i) - P(F\cap C_i)}\frac{P(C_i)}{P(C_{i+1})}. \tag{11}$$

Since, $x \sim^* y$ it follows that $P(F\cap C_{i-1}) + P(F\cap C_i) + P(F\cap C_{i+1}) = P(E\cap C_{i-1}) + P(E\cap C_i) + P(E\cap C_{i+1})$. This, with equations (10) and (11), imply that $\delta_i P(C_{i-1}) + \eta_i P(C_{i+1}) = P(C_i)$ as required by equation (4).

To prove the uniqueness it is enough to present two acts f and g in $\mathcal{F}$, such that if $P^f = P^g$ then necessarily $P^f = P$. Chose a full act f in $\mathcal{F}$. Then, for each E, $P^f(E) = \sum_i p_i P^f(E \mid C_i^f)$, for some consequential probability vector (p_i). By Proposition 14, $P^f(E) = \sum_i p_i P(E \mid C_i^f)$. To show that $P^f = P$, we need to show that $p_i = P(C_i^f)$.

Let g be an act such that for each i and j, $P(C_i^f \cap C_j^g) > 0$. Then f is a full act and therefore $f \in \mathcal{F}$. Now, $P^g(E) = \sum_i q_i P(E \mid C_i^g)$, for some consequential probability vector (q_i). Suppose $P^f = P^g$. Then for each j and k, $\sum_i p_i P^f(C_j^f \cap C_k^g \mid C_i^f) = \sum_i q_i P^g(C_j^f \cap C_k^g \mid C_i^g)$. All the added terms in each of these two sums are 0 but one. We conclude that $p_j P^f(C_j^f \cap C_k^g \mid C_j^f) = q_k P^g(C_j^f \cap C_k^g \mid C_k^g)$. These n^2 equations plus the equations $\sum_i p_i = \sum_i q_i = 1$ as equations in the $2n$ variables (p_i) and (q_i) are independent and hence can have at most one solution. Since $p_i = P(C_i^f)$ and $q_i = P(C_i^g)$ solve these equations, they are the unique solution. □

We say that the full acts f and g *overlap* if for each i, $C_i^f \cap C_i^g$ is a non-null event.

Lemma 8. *For all full acts f and g there exists a full act h such that f and h overlap and g and h overlap.*

Proof. By the non-atomicity of the measures P we can choose for each i and j, a partition of $C_i^f \cap C_j^g$ into two events $E_{i,j}$ and $F_{i,j}$ of equal probability. Let $E_i = \cup_j E_{i,j}$ and $F_j = \cup_i F_{i,j}$. Since $P(E_i) = (1/2)P(C_i^f) > 0$, E_i is non-null, and similarly F_j is non-null. Thus the act h defined by $C_k^h = E_k \cup F_k$ is in $\mathcal{F}$. Hence, for each i, $E_i \subseteq C_i^f \cap C_i^h$, f and h overlap. Similarly, g and h overlap. □

Proposition 16. *For all full acts f, g and i and j, $C_i^f \succsim^f C_j^f$ if and only if $C_i^g \succsim^g C_j^g$.*

Proof. Assume first that f and g overlap. By axiom Com 5 of Consequence Events, $C_i^f \succsim^f C_j^f$ if and only if $C_i^f \cap C_i^g \succsim^f C_j^f \cap C_j^g$. By axiom Act 2 of Common Desirability this relation holds if and only if $C_i^f \cap C_i^g \succsim^g C_j^f \cap C_j^g$, which, again, by axiom Com 5 of Common Consequences, holds if and only if $C_i^g \succsim C_j^g$. By Lemma 8, the claim holds also for non-overlapping acts. □

By Proposition 16 we can assume without lose of generality that for all full acts f,

$$C_n^f \succsim^f C_{n-1}^f \succsim^f \cdots \succsim^f C_1^f, \tag{12}$$

By Lemma 15 there exists a unique P that belongs to $\mathcal{P}(\succsim^f)$ for all $f \in \mathcal{F}$. For each such f there exists a unique utility u^f (determined up to a positive affine transformation) such that (P, u^f) represent $\succsim^f$. We now show that the same u serves for all f.

Proposition 17. *There exists a unique utility vector u which is determined up to a positive affine transformation, such that (P, u) represents $\succsim^f$ for each $f \in \mathcal{F}$.*

Proof. We assume first that f is a full act and that the desirability relations in equation (12) are strict. The argument for the general case is similar to the one used in the proof of Theorem 1.

We fix $f \in \mathcal{F}$ and omit all superscripts referring to this act. By Proposition 14, $P \in \mathcal{P}(\succsim^f)$ and thus, By equations (7), (10), and (11), the vector u is determined by,

$$\frac{\Delta u_{i+1}}{\Delta u_i} = \frac{\delta_i p_{i-1}}{\eta_i p_{i+1}} = \frac{P(F \cap C_{i-1}) - P(E \cap C_{i-1})}{P(E \cap C_{i+1}) - P(F \cap C_{i+1})}. \tag{13}$$

Suppose that f and g overlap. Then, it is possible to choose the events E and F in $\cup_i C_i^f \cap C_i^g$. For such events, $E \cap C_j^f = E \cap C_j^g$ for all j and similarly for F. Thus, when we compute the utility vector of g the right-hand side of equation (13) is the same for g and f. We conclude that for the same utility u, (P, u) represents both $\succsim^f$ and $\succsim^g$. By Lemma 8, (P, u) represent all $\succsim^f$ for all full f.

For acts f which are not full, the definition of overlapping acts cannot be used. In this case we say that a full act h *covers* an act f if the events $C_i^h \cap C_j^f$ is non-null for all i and all j such that C_j^f is non-null. The existence of a joint covering for any two acts can be easily proved similarly to Lemma 8. The rest of the proof easily follow.

□

5 An example

We discuss here the desirability relation $\succsim$ from Examples 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 3, 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 he 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:

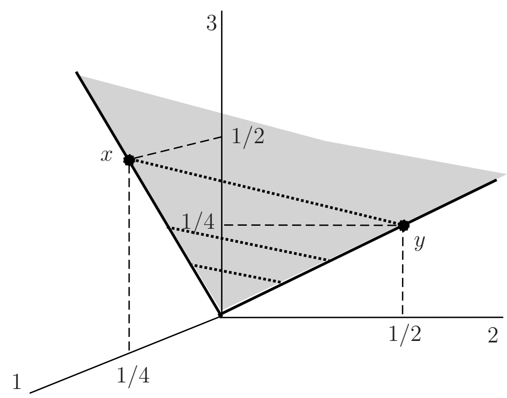
$$(1/3)\delta + (1/3)\eta = (1/3)(1), \tag{14}$$

$$(1/3)(0)\delta + (1/3)(1)\eta = (1/3)(1)(1/2). \tag{15}$$

Equation (14) reflects the preservation of probability. Since, the probability is kept fixed, equation (15) reflects that preservation of the *conditional* expected utility. Observe also, that these equations are the same for all x.

Equations (14) and (15) 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 (14) and (15) 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 Figure 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.



The cone generated by x and y consists of similar points. Each doted line consists of points which as well as being similar have also the same probability.

Figure 1: Similarity and same probability

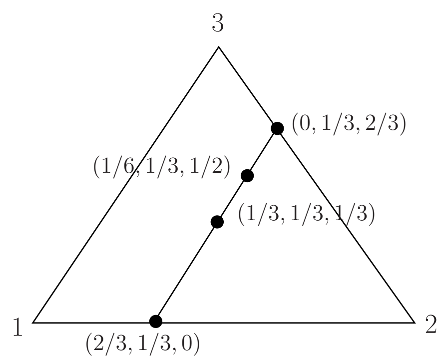
We now show the other pairs (Q, v) that represent the same relation $\succsim$. First, we know by Theorem 2 that $Q(\cdot \mid C_i) = P(\cdot \mid C_i)$ for each i. Thus, the projection of the Q-non-null events to $\mathbb{R}^3_+$ is the same as the projection of the P-non-null events. Also, since (P, 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 Com 7 of Persistency. 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 Figure 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:

$$q_1(1/2) + q_3(1/2) = q_3(1), \tag{16}$$

$$q_1 v_1(1/2) + q_3 v_3(1/2) = q_2 v_2(1). \tag{17}$$

The positive probabilities that solve (16) form an open interval of probabilities between $(2/3, 1/3, 0)$ and $(0, 1/3, 2/3)$ as in Figure 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.



The closer the point is to (0, 1/3, 2/3) the more optimistic it is.

Figure 2: The interval of consequential probabilities

Fixing q that solves (16) and solving for v in (17) 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. 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.

6 Concluding comments

A reader who even skimmed through the previous sections would be aware of some obvious or less obvious open questions and possible continuations of this work. We mention here just two of them.

First, the idea of comprehensive state spaces, as introduced by Aumann (1987), is one of the ideas that inspired our work. However, we have dealt only with comprehensive state spaces for a single individual. The next step should be the derivation of probability and utility in multi-agent comprehensive state spaces, and the study of rationality and equilibrium, as in Aumann (1987), in such spaces.

Second, we followed here the orthodoxy of the theory of decisions that considers binary choices as observable (in principle). Modern research in psychology and brain sciences redefines the concept of observability. It includes reportable statements, expressing for example desirability, but excludes utility and probability and gives up observable (in principle) binary choices of acts. Bridging between desirability as discussed here and this modern research calls for a different type of work.

References

Ahn, D. S. (2008), Ambiguity without a State Space, *The Review of Economic Studies*, 75, pp. 3-28.

Aumann, R. (1987), Correlated Equilibrium as an Expression of Bayesian Rationality, *Econometrica*, 55, pp. 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.

Chew, S. H. (1983), Axiomatic utility theories with the betweenness property, *Annals of Operations Research*, 19(1), 333–340.

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

Dekel, E. and B. L. Lipman and A. Rstichini(2001), Representing Preferences with a Unique Subjective State Space, *Econometrica*, 69(4), 891-934.

Jeffrey, R. (1965), *The Logic of Decision*, McGraw Hill, New York.

Jeffrey, R. (1983), *The Logic of Decision*, The University of Chicago Press, Chicago.

Jeffrey, R. (2004), *Subjective Probability: The Real Thing*, Cambridge University Press.

Gul, F. and W. Pesendorfer(2001), Temptation and Self-Control, *Econometrica*, 69(6), 1403-1435.

Kreps, D. M. (1979), A Representation Theorem for "Preference for Flexibility", *Econometrica*, 47(3), 565-577.

Kadane, J. B. and P. D. Larkey, (1982), Subjective probability and the theory of games, *Manage. Sci.*, 28, 113-120.

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

Ramsey, F. P. (1926), Truth and Probability in *Studies in Subjective Probability*, ed. H. Kyburg, Jr. and H. Smokler, Wiley, New York, 1964.

Savage, L. J. (1954), *The Foundations of Statistics*, Wiley, New York.

Telpaz, A. and D. J. Levy(2015), Consumers.Future Choices, *Journal of Marketing Research*, 52, 511-529.

von Neumann, J. and O. Morgenstern, (1953), *Theory of Games and Economic Behavior* , 3rd ed., Princeton University Press, Princeton.