Nationaløkonomisk Tidsskrift 130 (1992): 276-286

An Introduction to Decision Making When Uncertainty Is Not Just Risk

Ebbe Hendon, Hans Jørgen Jacobsen, Birgitte Sloth, and Torben Tranæs
Institute of Economics, University of Copenhagen

SUMMARY: *An uncertain and not just risky situation may be modeled using so-called belief functions assigning lower probabilities to events, that is, to sets of possible outcomes. In this paper we give an introduction to this concept and a first contribution to modeling decision making when there is uncertainty. We extend the von Neumann-Morgenstern expected utility theory to belief functions.*

1. Introduction

An economic *decision* is a choice among a set of possible *alternatives*. In some cases the alternatives may be equal to the *final outcomes*, but very often the final outcome also depends upon some *state of nature* which may be unknown when the decision is made. For example, a decision maker considers which share to buy in the stock market. The final outcome is what is earned on the transaction (in e.g. DKK), but this may depend upon future oil prices, on who will be the future manager of the firm, and other matters not known when the share is to be bought. Each decision can thus be viewed as a mapping assigning to each state of nature a particular outcome.

In the von Neumann-Morgenstern (1947) expected utility theory, the decision maker holds an exogenous probability distribution over the states of nature, and thereby, for each possible decision, a probability distribution over the final outcomes (a lottery). It is assumed that the individual has a preference relation ordering all lotteries, and that this preference relation fulfills some basic and appealing postulates – axioms. Then it is shown that the preference relation can be represented by a utility function fulfilling the *expected utility hypothesis*, i.e., the utility of any lottery is the expectation of the utilities of the final outcomes with respect to the probabilities.

In the von Neumann-Morgenstern theory the probabilities over states are exogenous[1]. However, in real life decision situations the present evidence consists of certain statements known to be true and from these the decision maker must make up what he

This paper reports part of the material in Hendon, Jacobsen, Sloth, and Tranæs (1991). We are very grateful to David Schmeidler for help and encouragement. Financial support from the Danish Social Sciences Research Council and Economic Department, Tel Aviv University is gratefully acknowledged.

1. This is what distinguishes it from the contributions of Savage (1954) and Anscombe-Aumann (1963), which are concerned with deriving subjective probabilities over states from preferences over decisions.

Nationaløkonomisk Tidsskrift 130 (1992): 276-286

An Introduction to Decision Making When Uncertainty Is Not Just Risk

Ebbe Hendon, Hans Jørgen Jacobsen, Birgitte Sloth, and Torben Tranæs
Institute of Economics, University of Copenhagen

SUMMARY: *An uncertain and not just risky situation may be modeled using so-called belief functions assigning lower probabilities to events, that is, to sets of possible outcomes. In this paper we give an introduction to this concept and a first contribution to modeling decision making when there is uncertainty. We extend the von Neumann-Morgenstern expected utility theory to belief functions.*

---

1. Introduction

An economic *decision* is a choice among a set of possible *alternatives*. In some cases the alternatives may be equal to the *final outcomes*, but very often the final outcome also depends upon some *state of nature* which may be unknown when the decision is made. For example, a decision maker considers which share to buy in the stock market. The final outcome is what is earned on the transaction (in e.g. DKK), but this may depend upon future oil prices, on who will be the future manager of the firm, and other matters not known when the share is to be bought. Each decision can thus be viewed as a mapping assigning to each state of nature a particular outcome.

In the von Neumann-Morgenstern (1947) expected utility theory, the decision maker holds an exogenous probability distribution over the states of nature, and thereby, for each possible decision, a probability distribution over the final outcomes (a lottery). It is assumed that the individual has a preference relation ordering all lotteries, and that this preference relation fulfills some basic and appealing postulates – axioms. Then it is shown that the preference relation can be represented by a utility function fulfilling the *expected utility hypothesis*, i.e., the utility of any lottery is the expectation of the utilities of the final outcomes with respect to the probabilities.

In the von Neumann-Morgenstern theory the probabilities over states are exogenous[1]. However, in real life decision situations the present evidence consists of certain statements known to be true and from these the decision maker must make up what he

---

This paper reports part of the material in Hendon, Jacobsen, Sloth, and Tranæs (1991). We are very grateful to David Schmeidler for help and encouragement. Financial support from the Danish Social Sciences Research Council and Economic Department, Tel Aviv University is gratefully acknowledged.

1. This is what distinguishes it from the contributions of Savage (1954) and Anscombe-Aumann (1963), which are concerned with deriving subjective probabilities over states from preferences over decisions.

believes about the plausibility of different states. Sometimes he is naturally led to a probability distribution over states, e.g. the roulette, but in other cases the information is not sufficiently rich for this. For instance, let there be three possible future managers r, s and t of the firm. No evidence speaks for any particular of them. It is known that r and s, but not t, graduated from the Harvard Business School. The majority of shareholders have confidence in candidates from Harvard, etc. In this and many other situations of economic interest, it seems highly unlikely that the decision maker, from the present evidence, should formulate his beliefs as precise as a probability distribution over states. To phrase it in classic terms: Uncertainty may be of a more fundamental nature than risk, that is, than what is captured by a probability distribution, as advocated by Knight and Keynes.

A prominent example is the so-called *Ellsberg paradox,* Ellsberg (1961). The objective, present evidence is this: There is an urn containing 150 balls; 50 of these are red, and 100 are either black or green; there is no further information. Individuals are first asked to choose between two bets, A and B, governed by a random draw of one ball. A: DKK 100 are won if the ball is red. B: DKK 100 are won if the ball is black. They are then given two other bets. C: DKK 100 if the ball is red *or* green. D: DKK 100 if the ball is black *or* green. A non-negligible proportion of decision makers strictly prefers A to B, and D to C. Assuming that bet I is strictly preferred to bet II if and only if bet I gives DKK 100 with strictly higher probability than bet II, these preferences are inconsistent with *any* additive probability over colors.

Generally, it seems relevant to consider decision environments, where an individual views each possible decision as giving rise to a mapping from states to outcomes but where the individual's assessment of which state will occur is uncertain in a fundamental way that cannot be captured by a probability distribution over the states, and where the consequence of each possible decision is therefore something more fundamentally uncertain than a probability distribution over outcomes. As the objects capturing this more basic uncertainty, we use a generalization of probability distributions namely *belief functions,* Shafer (1976), or *lower probability functions,* Dempster (1967). Belief functions are formally introduced and interpreted in Section 2 below.

The primitives of the theory presented here are the set of belief functions over a finite set of outcomes and a preference relation on this set. The task will be to derive a utility representation of this preference relation from axioms on it, thereby obtaining some first insight in the way decision makers may deal with uncertainty. Our theory is von Neumann-Morgenstern-like in the sense that the belief functions over states (outcomes) are exogenous. There is also a literature concerned with *deriving* subjective, non-additive probabilities over states, Schmeidler (1989), Gilboa (1987) and Wakker

(1986)[2]. However, the contribution closest to the present one is the work of Jaffray (1989).

The axioms we use here are as standard as possible in order not to confuse things. In fact, we are going to use the von Neumann-Morgenstern axioms and then add one which is supposed to be weak. The present paper is thus a first step in representing preferences over belief functions.

2. Belief functions and the problem considered

Consider a finite set of *outcomes;* $X = (x_1, ..., x_n)$, and denote by χ the set of all subsets of X. An element E of χ is referred to as an *event.*

A *probability measure* π assigns to each event E a number $\pi(E)$, the probability of E. A probability measure is thus a function $\pi: \chi \to [0,1]$, fulfilling:

$$\pi(\emptyset) = 0 \text{ and } \pi(X) = 1, \tag{1}$$

$$\pi(E \cup F) = \pi(E) + \pi(F) - \pi(E \cap F), \text{ for any two events } E \text{ and } F. \tag{2}$$

Property (2) is called *additivity.* The set of probability measures on X is denoted V^{Λ}. For convenience, write $\pi(x)$ for $\pi(\{x\})$ for all $x \in X$. From (2), $\pi(E) = \sum_{x \in E} \pi(x)$ for any E.

A belief function $v: \chi \to [0, 1]$, is a generalization of a probability measure obtained by weakening (2). The interpretation of a belief function is that it assigns to each event E, a lower bound $v(E)$, on the likelihood of E, see also Shafer (1976). The *weight of the evidence in support* of E, is $v(E)$, while the *plausibility* of E is $1-v(\bar{E})$, where $\bar{E}$ is the complement of E. So, a belief function embodies in this interpretation both a lower and a upper bound on the likelihood of each event, and in this sense it may contain uncertainty *additional to risk:* It assigns to each event not just a single number, the probability of that event, but an interval, the *range of possible probabilities* of the event.

As an example consider again the share/manager example of the introduction and assume that our decision maker, to the three possible states r, s, and t, associates the three rates of return x_1, x_2, and x_3 respectively. The fact that he has little faith in any particular of the potential future managers, but high faith in it being either r or s could, for instance, be expressed by the following belief function over outcomes: $v(\emptyset) = 0$, $v(\{x_1\}) = v(\{x_2\}) = v(\{x_3\}) = 0.1$, $v(\{x_1, x_3\}) = v(\{x_2, x_3\}) = 0.2$, $v(\{x_1, x_2\}) = 0.8$ $v(\{x_1, x_2, x_3\}) = 1$.

Which properties should be attributed to a belief function v? One suggestion is that for any two events E and F, $v(E \cup F) \geq v(E) + v(F) - v(E \cap F)$. The idea, which we

2. Alternative approaches are Gilboa and Schmeidler (1989) and Vind (1991).

will soon make precise, is that going to larger sets can only increase certainty. In the same spirit it could then be suggested that for any three events E, F and G, we should have $v(E \cup F \cup G) \geq v(E) + v(F) + v(G) - v(E \cap F) - v(E \cap G) - v(F \cap G) + v(E \cap F \cap G)$, etc. In Schafer (1976), the following restriction, which is a generalization of the above for two and three events to k events, is introduced as the defining property of a belief function:

$$v(E_1 \cup \ldots \cup E_k) \geq \sum_{I \subseteq \{1,\ldots,k\}} (-1)^{\#I+1} v(\bigcap_{i \in I} E_i), \text{ for any } k \text{ events } E_1,\ldots,E_k. \quad (3)$$

This property is called k-monotonicity. To understand the content of k-monotonicity note that (3) with *equality* is the usual inclusion-exclusion rule for probability measures which *follows* from (2) by induction. As already noted, a similar implication does not hold for *inequalities.* The condition of k-monotonicity must therefore be imposed in order to have an analog to the usual inclusion-exclusion rule.

We require k-monotonicity for all k, *i.e.,* $v : \chi \rightarrow [0,1]$ is a belief function if it satisfies (1), and (3) for any $k \geq 2$. Denote by V^B the set of all belief functions. Since probability measures are k-monotone for all k, we have $V^A \subseteq V^B$. One can simply think of a belief function v as the vector $(v(E))_{E \in \chi}$ *in* $\mathbb{R}^m$, $m := \#X$. For any two belief functions v_1, $v_2 \in V^A$, and $\alpha \in [0,1]$, the convex combination $v = \alpha v_1 + (1-\alpha) v_2$ is defined by: $v(E) = \alpha v_1(E) + (1-\alpha) v_2(E)$, for all $E \in \chi$. It is easy to verify that V^A and V^B are convex sets.

We have not yet *justified* the assumption of k-monotonicity for all $k \geq 2$. However, one very natural way to think of an uncertain environment is as given by a so-called *massfunction,* as in Shafer (1976). A function $m : \chi \rightarrow [0,1]$, is a mass function if $m(\emptyset) = 0$, and $\sum_{E \in \chi} m(E) = 1$. The interpretation is that for any event E, $m(E)$ is the weight of evidence in support of E which is additional to the weight already assigned to the proper subsets of E. The fact that m has non-negative values captures the idea that going to larger events can only increase certainty. Think of each possible decision as giving rise to a particular mass function m. The belief in an event F is naturally defined as: $v(F) := \sum_{E: E \subseteq F} m(E)$. Shafer (1976) shows that if v is defined from a mass function in this way, then v is a belief function, and conversely, any belief function is given by a uniquely determined mass function.

Returning to the share/manager problem we see that v can be derived from the mass function m, where $m(\{x_1\}) = m(\{x_2\}) = m(\{x_3\}) = m(\{x_1, x_2, x_3\}) = 0.1$, $m(\{x_1, x_2\}) = 0.6$, and $m(\{x_1, x_3\}) = m(\{x_2, x_3\}) = 0$. So most of the evidence weighs in favor of $\{x_1, x_2\}$ reflecting that the "Harvard argument" points to this set.

A particular class of belief functions is the *unit belief functions.* For any $E \in X$, $E \neq \emptyset$, define v_E by: $v_E(F) = 1$ if $E \subseteq F$, and $v_E(F) = 0$ otherwise. This says that the outcome will be an element in E for sure, and nothing else. Then v_E is indeed a belief function: Its mass function is simply $m(F) = 0$ for all $F \neq E$, and $m(E) = 1$.

Let m_v be the unique mass function defining v. Then for any F, $v(F) = \sum_{E:E \subseteq F} m_v(E)$ $= \sum_{E, \in X \setminus \{\emptyset\}} m_v(E) v_E(F)$, so v is a convex combination of the unit belief functions:

$$v \;\; = \sum_{E \in \chi \setminus \{\emptyset\}} m_v(E) v_E \tag{4}$$

The primitive of our theory is $(V^B, \succsim)$, the set of belief functions over X with a preference relation $\succsim$ on it, where $\succsim$ should be read "is at least as good as".[3] The problem is, from certain axioms on $\succsim$, to find a representation; a function $U : V^B \rightarrow \mathbb{R}$, such that for all $v, w \in V^B$: $v \succsim w \Leftrightarrow U(v) \geq U(w)$, where U has an intuitive interpretation.

From $\succsim$, define $\succ$ by: $v \succ w$ if $v \succsim w$ but not $w \succsim v$, and define $\sim$ by: $v \sim w$ if $v \succsim w$ and $w \succsim v$. Given $(V^B, \succsim)$ there will also be given preferences over the probability measures in V^A, in particular over the lotteries assigning probability 1 to one outcome, and zero to all other outcomes. We write $x_1 \succsim x_2$ if the lottery giving x_1 for sure is at least as good as the lottery giving x_2 for sure. Without loss of generality we assume that $x_1 \succsim x_2 \succsim .... \succsim x_n$.

A simple version of the so-called Mixture Set Theorem, Herstein and Milnor (1953), will be of great use. Let M be a convex subset of $\mathbb{R}^k$, $k \in \mathbb{N}$, and let $\succsim$ be a relation on M. Define $\succ$ and $\sim$ from $\succsim$ as above.

A1. (Weak order). $\succsim$ on M is complete and transitive.

A2. (Independence). For all $m_1, m_2, m_3 \in M$, and $\alpha \in]0,1]$: If $m_1 \succ m_2$, then $\alpha m_1 + (1-\alpha)m_3 \succ \alpha m_2 + (1-\alpha)m_3$.

A3. (Continuity). For all $m_1, m_2, m_3 \in M$, such that $m_1 \succ m_2 \succ m_3$ there are α, $\beta \in [0,1]$, such that: $\alpha m_1, +(1-\alpha)m_3 \succ m_2 \succ \beta m_1 + (1-\beta)m_3$.

MIXTURE SET THEOREM, *MST. The following two statements are equivalent:*

(i) $(M, \succsim)$ *fulfills* $A1 - A3$.

(ii) There is an affine function $U: M \rightarrow \mathbb{R}^4$, *such that U represents* $\succsim$. *Further, if* U' *is another affine representation of* $\succsim$, *then* $U' = aU + b$, *for some* $a, b \in \mathbb{R}$, $a > 0$.

Applying the *MST* to $(V^A, \succsim)$ yields the von Neumann-Morgenstern theory of expected utility: From affinity of U, one gets that for any $\pi \in V^A$, $U(\pi) = U(\sum_{i=1}^{n} \pi_i e_i$

3. Let a decision be a function $f : S \rightarrow X$, from a finite set of states to X. Further, let uncertainty with respect to states be described by the belief function η: $\mathcal{S} \rightarrow [0,1]$, where $\mathcal{S}$ is the set of subsets of S. Derive v: $\chi \rightarrow [0,1]$ in the natural way: $v(E) = \eta(f^{-1}(E))$ for all $E \in \chi$. Then v is a belief function on X. This justifies that we start right off with belief functions on X.

4. The function U is affine if $U(\alpha m_1 + (1-\alpha)m_2) = \alpha U(m_1) + (1-\alpha) U(m_2)$, *for all* $m_1, m_2 \in M$, $\alpha \in [0,1]$.

$= \sum_{i=1}^{n} \pi_i U(e_i)$, where e_i is the i'th unit vector, that is the lottery giving x_i for sure. Defining $u(x_i) := U(e_i)$, we have: $U(\pi) = \sum_{x \in X} u(x)\pi(x)$. The utility assigned to a lottery π, is the expected value, with respect to π, of the utilities assigned to the sure outcomes.

The core[5], $C(v)$, of a belief function v is the set of probability measures which do not contradict v:

$$C(v) := \{\pi \in V^A \mid \pi(E) \geq v(E) \text{ for all } E \in \chi\}. \tag{5}$$

From Shapley (1971), for any 2-monotone v, $C(v)$ is non-empty. So, any belief function has a non-empty core, and the core of a probability measure π is $\{\pi\}$.

A function p from $\{1, \ldots n\}$ onto X, gives a particular sequence of the outcomes; $p(i)$ is the outcome on place i, $p^{-1}(x)$ is the place given to x. The set of all $n!$ permutations is P. For any v in V^B, and any p in P, we define the probability measure $\pi^*_{p,v}$ by:

$$\pi^*_{p,v}(p(1)) := v(\{p(1)\}), \text{ and} \tag{6}$$

$$\pi^*_{p,v}(p(i)) := v(\bigcup_{j=1}^{i}\{p(j)\}) - v(\bigcup_{j=1}^{i-1}\{p(j)\}) \text{ for } i \in \{2,\ldots,n\}.$$

To illustrate consider the simple permutation $p'(i) = x_i$. Then $\pi^*_{p',v}$ is the probability measure constructed from v by first giving to x_1, the best outcome, the least probability that can be given according to v. The second best outcome x_2 is then given the least probability *it* can be given according to v and what has already been given to x_1, etc.

Shapley (1971) shows that for all 2-monotone v, the set of all vertices (corners) of $C(v)$ is $\{\pi^*_{p,v} \in V^A \mid p \in P\}$, and $C(v)$ is the convex hull of this set.

An example will clarify these concepts. Consider again the belief function of the share/manager example. Its core is illustrated in Figure 1. The triangle represents the set of all probability distributions over $\{x_1, x_2, x_3\}$, with $\pi(x_1)$ measured along the $\pi(x_1)$–axis and so forth. From $v(\{x_3\}) = 0.1$, we obtain from (5) the restriction $\pi(x_3) \geq 0.1$. From $v(\{x_1, x_2\}) = 0.8$, $\pi(x_1) + \pi(x_2) \geq 0.8$, and hence $\pi(x_3) \leq 0.2$. Continuing like this the core of v appears as the shaded area. The points $\pi^*_{p,v}$ are easily computed. For instance, for $p = (x_2, x_3, x_1)$, it follows from (6), that $\pi^*_{p,v}(x_2) = 0.1$, then $\pi^*_{p,v}(x_3) = 0.2 - 0.1 = 0.1$, and finally $\pi^*_{p,v}(x_1) = 1 - 0.2 = 0.8$, so $\pi^*_{p,v} = (0.8, 0.1, 0.1)$, which is one of the vertices of the core. For this particular v, there are different permutations giving rise to the same $\pi^*_{p,v}$. With three outcomes, the core may have up to six vertices.

5. The core is a concept from cooperative game theory. Those familiar with this branch will realize that a belief function $v: \chi \rightarrow [0,1]$, has exactly the structure of a *characteristic function* of a game with side payment, where X would then be the set of players, and χ the set of coalitions.

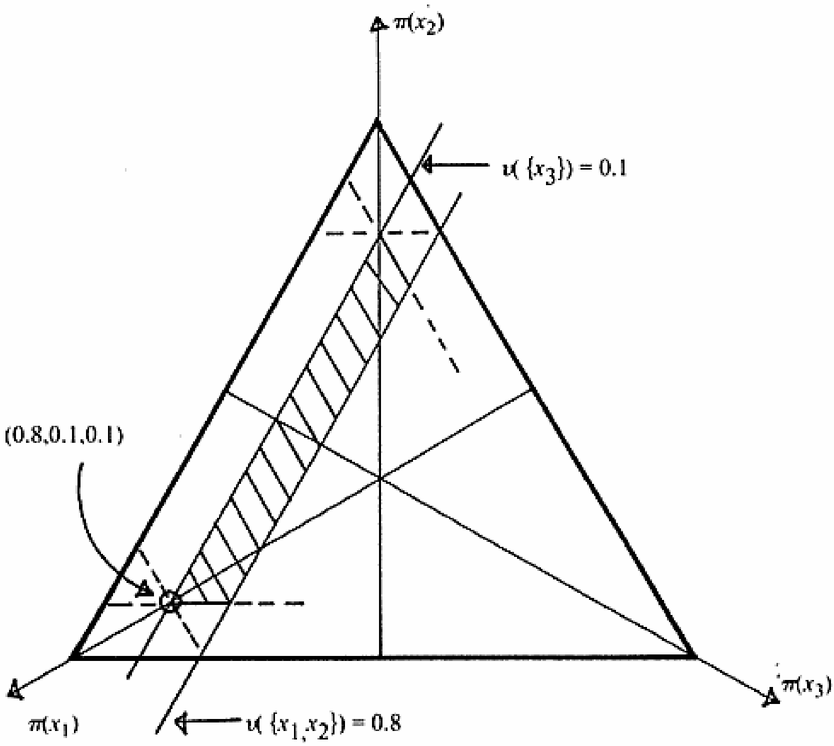

Figure 1. The core in the share/manager example

3. Representation theorems: The nature of $(V^B, \succsim)$

The set of belief functions is convex, so assuming A1-A3 for $(V^B, \succsim)$ would, directly from the *MST*, imply the existence of an affine representing function $U : V^B \rightarrow \mathbb{R}$ (and vice versa). Using (4) and affinity, U can be written: $U(v) = U(\sum_{E \in X \setminus \emptyset} m_v(E)\, v_E) = \sum_{E \in X \setminus \emptyset} m_v(E) U(v_E)$. This observation is parallel to the von Neumann-Morgenstern representation for probability measures, where now $U(v_E)$ is the counterpart of $u(x)$. In so far as one has good intuition for $U(v_E)$ – like one has in the von Neumann-Morgenstern theory for $u(x)$ – we have now an interesting representation. However, we do not think that the utility of ending in set E for sure but knowing nothing else, is as interpretable as the utility of getting outcome x for sure. Further, assuming A1-A3 for $(V^B, \succsim)$ implies assuming them for $(V^A, \succsim)$, so preferences over lotteries can still be represented using a von Neumann-Morgenstern utility on X. We would like to have the utility function U on V^B expressed in terms of u. To obtain this we must add an axiome:[6]

6. Jaffray (1989) also builds on the above observation. It is when it comes to "the additional axiom" that the present paper differs from that of Jaffray.

A4. (Non-extreme attitude towards uncertainty). For any $E \in \chi \setminus \{\emptyset\}$, there are $\underline{\pi}_E, \bar{\pi}_E \in V^A$, such that $\underline{\pi}_E(E) = \bar{\pi}_E(E) = 1$, and $\bar{\pi}_E \succsim v_E \succsim \underline{\pi}_E$.

We consider A4 to be very weak. First note that A4 is an axiom *only* on the unit belief functions. A unit belief function v_E, contains the information that the outcome will be in E for sure and nothing else. The set of probability measures giving something in E for sure is $\{\pi \in V^A \mid \pi(E) = \sum_{x \in E} \pi(x) = 1\}$, which is exactly $C(v_E)$. The elements of this set are ranked by $\succsim$. The content of A4 is just that v_E itself is no better than the very best element and no worse than the very worst element of this set. In particular, since A4 is made only on unit belief functions, $\underline{\pi}_E$ and $\bar{\pi}_E$ may be the worst and the best *pure* outcome of E that is, we could have $\underline{\pi}_E$, $\bar{\pi}_E \in E$, and $\bar{\pi}_E \succsim x \succsim \underline{\pi}_E$ for all $x \in E$.

Theorem 1 below is a rather easy consequence of the *MST* and A4. For formal proofs the reader is referred to Hendon, Jacobsen, Sloth, and Tranæs (1991).

THEOREM 1. *Let $(V^B, \succsim)$ be given. The following two statements are equivalent:*

(i) $(V^B, \succsim)$ fulfills A1-A4.

(ii) There is a function $u: X \to \mathbb{R}$, and for each $v \in V^B$, there is a probability measure $\pi_v \in C(v)$, where $\pi_.$ considered as a function from V^B to V^A is affine, such that $U: V^B \to \mathbb{R}$, defined by $U(v) := \sum_{x \in X} u(x)\pi_v(x)$ represents $\succsim$. Further, the function u is unique up to a positive affine transformation.

Theorem 1 is a qualitative statement on the nature of $(V^B, \succsim)$. Given A1-A4, preferences over the complicated objects, belief functions, are *as if* each belief function is identified with a lottery in its core, from which the expected von Neumann-Morgenstern utility in the usual sense is computed, and the so defined numbers rank all the belief functions in preference. Conversely, if the preference has such a representation, A1-A4 are fulfilled. However, Theorem 1 says little about *which* element of $C(v)$, v is identified with. We only know that the function $\pi_.$ from V^B *to* V^A is affine, and thereby continuous, in v. It is, of course, warranted to say more about the variation of π_v with v. Since Theorem 1 is an if-and-only-if-theorem, we know that this will require stronger axioms. We introduce now an axiom of *concistency* which is stronger than A4.

Consider the belief function v_X which contains absolutely no information: It says that the outcome will be in X and nothing else. Assume that A1-A2 are fulfilled. Then there is some subset $\Pi_X \subseteq V^A$, such that $v_X \sim \pi$ for all $\pi \in \Pi_X$. A natural *consistency requirement* will now be that there is some $\pi_X \in \Pi_X$, with $\pi_X(x) > 0$ for all x, such that for each subset E, $v_E \sim \pi_E \in V^A$, where π_E is derived from π_X by Bayes' rule, i.e.,

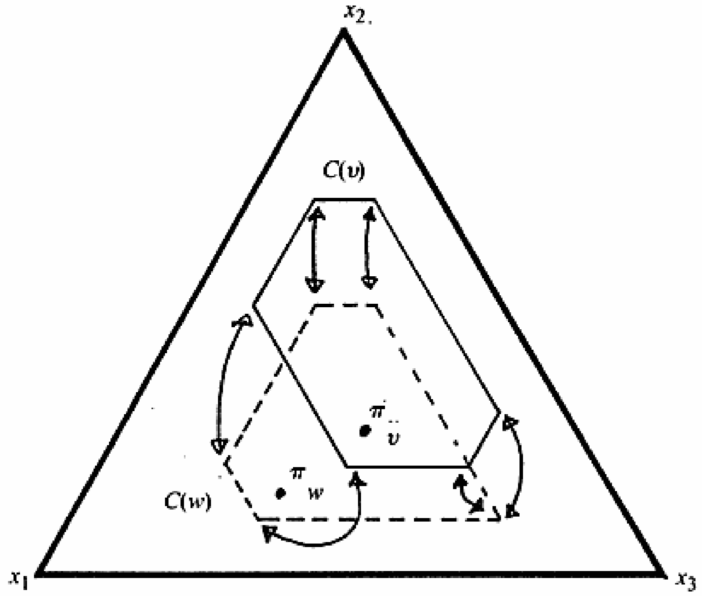

Figure 2.

$\pi_E(x) = \pi_X(x)/\pi_X(E)$ for $x \in E$.[7] This is a requirement of internal consistency of the decision maker's preference over *unit* belief functions. As the appropriate consistency concept we use Bayes' rule which is possible since the requirement is only concerned with unit belief functions. Note that Bayes' rule is *not* used for updating the uncertain objects, the belief functions; ist is *only* used as a way of expressing consistency of preferences.

A5. (Consistency). There is $\pi_X \in V^A$ with $\pi_X(x) > 0$ for all $x \in X$, such that $v_X \sim \pi_X$, and for each $E \in \chi\backslash\{\emptyset\}$, $v_E \sim \pi_E$, where π_E is given by $\pi_E(x) := \pi_X(x)/\pi_X(E)$, for all $x \in E$, and $\pi_E(x) = 0$ for $x \notin E$.

It is obvious that A5 is a special case of A4: Just use the π_E of A5 as both $\underline{\pi}_E$ and $\bar{\pi}_E$ in A4. To prove Theorem 2 below, basically what is needed is Theorem 1 above and a result from cooperative game theory on *Shapley values* (see e.g. Kalai and Samet (1988)).

7. The assumption *is* restrictive, and is made for simplicity. Axiom A5 and Theorem 2 can be generalised to allow for $\pi_X(x) = 0$ for some x by using the notion of a weight system, see Hendon, Jacobsen, Sloth, and Tranæs (1991).

THEOREM 2. *Let $(V^B, \succsim)$ be given. The following two statements are equivalent:*
(i) $(V^B, \succsim)$ *fulfills* A1-A3 and A5.
(ii) There is a function $u: X \to \mathbb{R}$, and there is $\pi \in V^A$, where $\pi(x) > 0$ for all $x \in X$, such that if for each $p \in P$:

$$\alpha_p := \prod_{j=1}^{n} [\ \pi(p(j)) / \sum_{k=1}^{j} \pi(p(k))\ ],$$

the function: $U : V^B \to \mathbb{R}$ defined by $U(v) := \sum_{x\in X}\ u(x)\ [\ \sum_{p\in P}\ \alpha_p \pi^{*}_{p,v}]\ (x)$ represents $\succsim$. Further, the function u is unique up to positive affine transformation.

Note that for the α_p's of *(ii)*, $\alpha_p \in [0,1]$ for all $p \in P$, and $\sum_{p\in P} \alpha_p = 1$. So, Theorem 2 gives a high degree of regularity in the choice of π_v from Theorem 1. A person fulfilling A1-A3 and A5 (and thereby A4) behaves as if he assigns to each belief function v a number, which is the expected von Neumann-Morgenstern utility of a π_v in the core of v, where π_v is computed as a weighted average of all the vertices $\pi^{*}_{p,v}$ of the core of v, and *the weights α_p are independent of v.* The regularity comes from the last part: The weighted avarage used is the same for all belief functions.

As an illustration consider Figure 2, where the cores of two different belief functions v and w have been drawn, fully and dotted, respectively. The vertices of the two cores are connected two by two by arrows. Connected vertices are given the same weights in the computations of π_v and π_w. The lower, right vertex of *C(w)* is given weight equal to the sum of the weights of the two lower, right vertices of *C(v)*. If, for instance, π_v is situated close to the vertex arising from the requirements given by $v(\{x_2\})$ and $v(\{x_1, x_2\})$ then π_w is also close to the corresponding vertex.

With a representation like that of Theorem 2 the Ellsberg paradox need no longer be a paradox. From the objective evidence, the belief function over colors (states) would naturally be: $\eta(\{red\}) = 1/3$, $\eta(\{black\}) = \eta(\{green\}) = 0$, $\eta(\{red,\ black\}) = \eta(\{red,\ green\}) = 1/3$, $\eta(\{black,\ green\}) = 2/3$. The derived belief functions over outcomes are then for each alternative: *A:* $v^a(100) = 1/3$, $v^a(0) = 2/3$; *B*: $v^b(100) = 0$, $v^b(0) = 1/3$; *C*: $v^c(100) = 1/3$, $v^c(0) = 0$; *D:* $v^d(100) = 2/3$, $v^d(0) = 1/3$. The cores of these can be expressed in terms of the probability of winning DKK 100: $C(v^a) = \{1/3\}$, $C(v^b) = \{\pi \mid \pi \leq 2/3\}$, $C(v^c) = \{\pi \mid \pi \geq 1/3\}$, $C(v^d) = \{2/3\}$. To each core the decision maker associates a specific probability measure. Within the axioms of Theorem 2, this could very well be close to the minimal possible probability of winning DKK 100 (this corresponds to choosing $\alpha_{p'}$, where $p' = (0,100)$, close to 1, which is obtained by letting $\pi_X(0)$ being close to 1 and $\pi_X(100)$ close to zero), i.e., close to $\pi_{va} = 1/3$, $\pi_{vb} = 0$, $\pi_{vc} = 1/3$, $\pi_{vd} = 2/3$. Assuming that $u(100) > u(0)$, calculation of expected utility with respect to these probabilities gives $A \succ B$ and $D \succ C$.

References

Anscombe, F. J. and R. J. Aumann. 1963. A Definition of Subjective Probability. *The Annals of Mathematical Statistics,* 34, 199-205.

Dempster, A. P. 1967. Upper and Lower Probabilities Induced by a Multivalued Mapping. *The Annals of Mathematical Statistics,* 38, 325-339.

Ellsberg, D. 1961. Risk, Ambiguity, and the Savage Axioms. *Quarterly Journal of Economics,* 75, 643-669.

Gilboa, I. 1967. Expected Utility with Purely Subjective Non-Additive Probabilities. *Journal of Mathematical Economics,* 16, 65-88.

Gilboa, I. and D. Schmeidler. 1989. Maxmin Expected Utility with Non-Unique Prior. *Journal of Mathematical Economics,* 18, 141-153.

Hendon, E., H. J. Jacobsen, B. Sloth and T. Tranæs. 1991. Expected Utility under Uncerntainty. Manuscript. Universitet of Copenhagen.

Herstein, I. N. and J. Milnor. 1953. An Axiomatic Approach to Measureable Utility. *Econometrica,* 21, 291-297.

Jaffray, J. Y. 1989. Application of Linear Utility Theory to Belief Functions. *Operations Research Letters* 8, 107-112.

Kalai, E. and D. Samet. 1988. Weighted Shapley Values. In A. E. Roth (ed.): *The Shapley Value,* Cambridge.

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

Schmeidler, D. 1989. Subjective Probability and Expected Utility without Additivity, *Econometrica,* 57, 571-587.

Shafer, G. 1976. *A Mathematical Theory of Evidence.* Princeton.

Shapley, L. S. 1971. Cores of Convex Games. *International Journal of Game Theory,* 1, 11-26.

Vind, K. 1991. Independent Preferences. *Journal of Mathematical Economics,* 20, 119-135.

von Neumann, J. and O. Morgenstern. 1947. *Theory of Games and Economic Behavior,* 2nd. ed. Princeton.

Wakker, P. P. 1986. *Representations of Choice Situations.* PhD-thesis, University of Tilburg.