No. 803

Decomposable Choice under Uncertainty

by

Simon Grant,Atsushi Kajii and Ben Polak

December 1998

December 8, 1998

Decomposable Choice under Uncertainty.*

Abstract

Savage motivated his Sure Thing Principle by arguing that, whenever an act would be preferred if an event obtains and preferred if that event did not obtain, then it should be preferred overall. The idea that it should be possible to decompose and recompose decision problems in this way has normative appeal. We show, however, that it does not require the full separability across events implicit in Savage's axiom. We formulate a weaker axiom that suffices for decomposability, and show that this implies an implicit additive representation. Our decomposability property makes local necessary conditions for optimality, globally sufficient. Thus, it is useful in computing optimal acts. It also enables Nash behavior in games of incomplete information to be decentralized to the agent-normal form. None of these results rely on probabilistic sophistication; indeed, our axiom is consistent with the Ellsberg paradox. If we assume probabilistic sophistication, however, then the axiom holds if and only if the agent's induced preferences over lotteries satisfy betweenness.

JEL Classification: D80, D81

Keywords: sure-thing principle, decomposability, uncertainty, computation, agent-normal form games, non-expected utility, betweenness.

Simon Grant
Department of Economics
Faculty of Economics and Commerce
Australian National University

Atsushi Kajii
Institute of Policy and Planning Sciences
University of Tsukuba

Ben Polak
Economics Department
Yale University

*An earlier version of this paper (October 1997) appeared under the title "Weakening the Sure-Thing Principle: Decomposable Choice under Uncertainty". We thank Didier Dubois, Larry Epstein, Chew Soo Hong, Mamoru Kaneko, Peter Klibanoff, François Ortalo-Magné, David Pearce, and Peter Wakker. Parts of this research was undertaken when Kajii was visiting ANU and Yale and when Polak was visiting LSE. We thank these institutions for their hospitality.

1 Savage and Decomposing Choice Problems

In Savage's axiomatization of subjective expected utility theory under uncertainty, his second postulate P2, often referred to as the 'sure-thing principle', is the analogue of the familiar independence axiom from standard expected utility theory under objective risk. Thus, theoretical and empirical criticisms of the separability assumptions implicit in the independence axiom apply equally to P2.[1] Savage originally motivated the sure thing principle, however, with an example that contains no explicit mention of separability.

> "A businessman contemplates buying a certain piece of property. He considers the outcome of the next presidential election relevant to the attractiveness of the purchase. So, to clarify the matter for himself, he asks whether he would buy if he knew that the Republican candidate were going to win, and decides that he would do so. Similarly, he considers whether he would buy if he knew that the Democratic candidate were going to win, and again finds that he would so. Seeing that he would buy in either event, he decides that he should buy, even though he does not know which event obtains, or will obtain, as we would ordinarily say." (Savage, 1972, p.21)

Decisions like buying property are difficult in part because each choice can give very different outcomes in different states of the world. The businessman's thought-experiment, breaking each alternative into component parts and comparing event by event, is a common way to simplify such problems. For this method to work, such decomposition and recomposition must yield the correct final decision: that is, if buying the property is preferred in all contingencies then it should be preferred overall. As Savage argued, this seems both a plausible and useful restriction to place on preferences: "except possibly for the assumption of simple ordering, I know of no other extralogical principle governing decisions that finds such ready acceptance."

Savage used the appeal of this idea to justify his second postulate. Indeed, it was this idea, rather than P2 itself, that Savage first referred to as the sure thing principle. In this paper, however, we argue that one can accept the idea of Savage's story while rejecting the separability assumptions implicit in Savage's postulate P2. Specifically, we propose a weaker rule — a weak decomposability principle — that allows the kind of decomposition and recomposition of decision problems used in the property-buying example, but that does not imply full separability across events.

[1] Machina (1983) provides a survey of the wide range of experimental violations of expected utility that have been observed. For the analogy between the Sure Thing Principle and Independence, see, for example, Machina & Schmeidler (1992). Schlee (1997) demonstrates, however, that the Sure Thing Principle plays a somewhat different role in ensuring dynamic consistency in decision making under uncertainty than does the Independence axiom in decision making under risk.

How then does Savage get from the nice property-buying story to the controversial separability implications of P2? The answer lies in his formalization. In Savage's framework, acts are functions from states of the world to outcomes. We can think of buying a property as an act, g, yielding different final outcomes depending on the state of the world. For the purpose of the example, not buying the property is also considered an act. Since it is hard to envisage an act defined negatively, let us interpret the property in question as a house and define the act f as 'not to buy and to stay put in the old residence'. Paraphrasing Savage, we can write the idea of decomposability as: if a person prefers an act g to another act f, either knowing that the event B obtains, or knowing that the complement of the event B obtains, then she should prefer g to f.

So far, so good. For Savage, however, preferences are only defined on the set of acts, and acts are functions from *all* states to outcomes. Objects like 'g if the event B obtains' are, at best, subacts. To formalize the statement that g is preferred to f if B obtains, Savage needs to extend these subacts over the whole domain. This was his next step.

> "What technical interpretation can be attached to the idea that g would be preferred to f, if B were known to obtain. Under any reasonable interpretation, the matter would seem not to depend on the values f and g assume at states outside of B. There is, then, no loss of generality in supposing that f and g agree with each other except in B; ... The first part of the sure-thing principle can now be interpreted thus: If after being modified so as to agree with one another outside of B, g is preferred to f; then g would be preferred to f if B were known."(Ibid. p.22)

This "technical interpretation" is less innocent than it at first appears. Consider, for example, the three acts, 'to buy the house', 'not to buy the house (and to stay put)', and 'to emigrate to Japan'. Savage interprets the statement that 'the agent would prefer to buy over not to buy if he knew the Republican would win' to mean that the agent prefers the act, say, 'buy if the Republican wins, and emigrate to Japan otherwise' to the act 'do not buy (and stay put) if the Republican wins, but still emigrate to Japan otherwise'. Moreover, the same preference must obtain if we change 'emigrate to Japan' to any other activity, regardless of its outcomes. But, it is precisely this separability — that preferences on the event B do not depend on what happens off B — that has been challenged on experimental and introspective grounds. For example, our hero might anticipate that his attitude toward owning and living in the house might be quite different if he knew that it came to him in place of a life in Japan. Compared to emigration, the house might now seem small and confining. Alternatively, compared to typical Japanese real estate, the house might now seem large and more attractive. Such thinking contradicts the separability implicit in Savage's particular "technical interpretation" of the idea in the story. But, Savage's technical interpretation is not necessary to capture this idea.

How else can one reasonably interpret the statement that "the agent would prefer to buy the house if he knew the Republican would win"? We take it simply to mean that the agent prefers the act 'to buy the house if the Republican wins, and not to buy otherwise' to the act 'not to buy the house regardless of the election results'. Under this interpretation, the statement has no implications about preferences among elaborate acts involving trips to Japan in the event of a Democratic victory: it only concerns preferences over acts constructed from the original acts 'to buy' or 'not to buy' the house. An axiom based on this interpretation, however, is sufficient to carry out the decomposition and recomposition used in the businessman's thought-experiment. In considering whether or not to buy the house, the businessman first compares the act 'to buy if the Republican wins but not otherwise' with the act 'not to buy regardless of the election'. He then compares the act 'to buy if the Democrat wins but not otherwise' with the act 'not to buy regardless'. Our axiom simply says that, if in both comparisons, the former is better than the latter, then he should prefer the act 'to buy regardless' to the act 'not to buy regardless'. That is, our axiom directly formalizes the idea in Savage's example under our less restrictive interpretation of statements involving preferences over sub-acts. We call this axiom *weak decomposablity.* The purpose of this paper is to explore the implications of this axiom: what representations are consistent with decomposability, and how it relates to other properties of preference relations.

Although the context is different, this paper is closely related to that of Gul & Lantto (1990). Their general approach to non-expected utility theory was "to consider possible applications ..., to formulate additional normative restrictions to the theory based on such applications [and] incorporate those restrictions [in]to the theory" (1990, p.173). Gul & Lantto's application is to dynamic choice under risk. The new normative restrictions they suggest are based on weakening consequentialism while keeping other desirable features. Our application is to static choice under uncertainty. Our new normative restriction is based on weakening the sure thing principle while still allowing choice to be decomposed. One of Gul & Lantto's normative restrictions, "dynamic programming solvability", can itself be thought of as a decomposability principle. Indeed, we show below that weak decomposability is equivalent to at least a version of dynamic programming solvability translated to a Savage framework.[2]

Skiadas (1997a,b) offers a different way to decompose a decision problem. In his framework, loosely speaking,[3] a decision maker is endowed with a richer set of preferences so that

[2] Gul & Lantto show that dynamic programming solvability is equivalent to two other weakenings of consequentialism in the context of dynamic choice, but we do not know how to translate these into a Savage framework.

[3] One reason this is loose is that Skiadis also does not define outcomes as Savage does.

not only can she compare any two acts, f and g, unconditionally, but she can also express a preference between some event E occurring given that she chose f and E occurring given that she chose g. His decomposability property ("strict coherence") says that if the agent prefers E given she chose f to E given she chose g, and the same is true of E complement, then she prefers f to g. Unlike our decomposability property, Skiadas's does not restrict preferences with respect to the substitution of one sub-act for another within a parent act. Both Savage and our restrictions are within a single preference relation, whereas Skiadas's applies across sets of preference relations. One could interpret this as a difference in the set of choices by the agent that the analyst can observe.

Section 2 introduces weak decomposability and discusses its properties. Section 3 shows that weak decomposability ensures that preferences admit at least an implicit additive representation. Section 4 shows how weak decomposability is related to: an algorithmic approach to find optimal plans; decentralized choice in agent-normal form games; and Gul & Lantto's "dynamic programming solvability". Up to this point, we do not assume probabilistic sophistication. Indeed in Section 5 we give an example of a preference relation that rationalizes choice behavior compatible with the Ellsberg paradox (hence it cannot be probabilistically sophisticated), while still satisfying weak decomposability and all the Savage postulates except P2. However, if the agent is in fact probabilistically sophisticated, then weak decomposability is equivalent to Chew's (1983, 1989) and Dekel's (1986) betweenness property. Thus, just as Savage's original sure thing property forms part of the axiomatization of subjective expected utility theory, so weak decomposability forms part of an axiomatization of subjective betweenness theory. Unless stated otherwise, proofs of all observations and propositions are to be found in the appendix.

2 The Weak Decomposability Principle

Denote by $\mathcal{S} = \{\ldots, s, \ldots\}$ a set of states, $\mathcal{E} = \{\ldots, A, B, \ldots, E, \ldots\}$ the set of events which is a given σ-field on $\mathcal{S}$, and $\mathcal{X} = \{\ldots, x, y, z, \ldots\}$ a set of outcomes or consequences. An act is a (measurable) function $f : \mathcal{S} \rightarrow \mathcal{X}$. Let $f(\mathcal{S}) = \{f(s) \,|\, s \in \mathcal{S}\}$ be the outcome set associated with the act f, and let $\mathcal{F} = \{\ldots, f, g, h, \ldots\}$ denote the set of acts on $\mathcal{S}$ with finite outcome sets. We will abuse notation and use x to denote both the outcome x in $\mathcal{X}$ and the constant act $f(\mathcal{S}) = \{x\}$. Let $\succeq$ be a binary relation over ordered pairs of acts in $\mathcal{F}$, representing the individual's preferences. Let $\succ$ and $\sim$ correspond to strict preference and indifference, respectively. Denote by $\succeq_{\mathcal{X}}$ the relation over ordered pairs of outcomes obtained from $\succeq$ for constant acts (that is, $x \succeq_{\mathcal{X}} y$ if and only if $x \succeq y$).

The following notation to describe an act will be convenient. For an event E in $\mathcal{E}$, and any two acts f and g in $\mathcal{F}$, let $f_E g$ be the act which gives, for each state s, the outcome $f(s)$ if s is in E and the outcome $g(s)$ if s is in the complement of E (denoted $\mathcal{S}\backslash E$). In general, for any finite partition $\{A_1, \ldots, A_n\}$ of $\mathcal{S}$ and any set of n acts $\{h^1, \ldots, h^n\}$, let $h^1_{A_1} h^2_{A_2} \ldots h^{n-1}_{A_{n-1}} h^n$ be the act that yields $h^i(s)$ if s is in A_i. Using this notation we can define the set of *null* events, $\mathcal{N} \subset \mathcal{E}$, as follows: $E \in \mathcal{N}$ if and only if, for all acts f, g and h $\in \mathcal{F}$, $f_E h \sim g_E h$.

Savage's first six postulates, together with Machina & Schmeidler's (1992) names for them, are as follows:

P1 (Ordering): The preference relation $\succeq$ is complete, reflexive and transitive.

P2 (Sure Thing Principle): For all events E and acts f, g, h and h', if $f_E h \succeq g_E h$ then $f_E h' \succeq g_E h'$.

P3 (Eventwise Monotonicity): For all non-null events E, pairs of outcomes x and y, and acts h, $x_E h \succeq y_E h$ if and only if $x \succeq_{\mathcal{X}} y$.

P4 (Weak Comparative Probability): For all events A and B, and outcomes $x^* \succ_{\mathcal{X}} x$, and $y^* \succ_{\mathcal{X}} y$, if $x^*_A x \succeq x^*_B x$ then $y^*_A y \succeq y^*_B y$.

P5 (Nondegeneracy): There exist outcomes x and y such that $x \succ y$.

P6 (Small Event Continuity) For any pair of acts $f \succ g$ and outcome x, there exists a finite set of events $\{A_1, \ldots, A_K\}$ forming a partition $\mathcal{P}$ of $\mathcal{S}$, such that for all A in $\mathcal{P}$, $x_A f \succ g$ and $f \succ x_A g$.

In what follows, we will always assume (stated or otherwise) that preferences satisfy Savage's ordering assumption, P1. We will also always assume P5, not because it is necessary for the results but because the problem is trivial without it. We will not assume the other postulates unless explicitly stated.

For the last result of this section and for some of the results in section 4 we will use the following strengthening of **P6** that is implied by Savage's six postulates.

P6* (Event Continuity) For all acts f, g, and h in $\mathcal{F}$, all outcomes x, z in $\mathcal{X}$, and all events A in $\mathcal{E}$, if $f_A h \succ g_A h$ and $x \succeq_{\mathcal{X}} y \succeq_{\mathcal{X}} z$ for all y in $f(A) \cup g(A)$, then there is an event $E \subseteq A$ such that $x_E g_{A \backslash E} h \sim f_A h$ and an event $E^* \subseteq A$, such that $z_{E^*} f_{A \backslash E^*} h \sim g_A h$.[4]

[4] Notice that A can be neither empty nor null since $f_A h \succ g_A h$. We think it is not known whether P6* is implied by P1, P3 and P6, even when $\mathcal{E}$ is the set of all subsets of $\mathcal{S}$.

In Section 5, we use Epstein & Le Breton (1993) strengthening of Savage's weak comparative probability axiom, P4.

P4^c (Conditional Weak Comparative Probability): For all events T, A and B, $A \cup B \subseteq T$, outcomes x^*, x, y^*, and y, and acts g, if $x^*_T g \succeq x_T g$ and $y^*_T g \succeq y_T g$, then $x^*_A x_{T\setminus A} g \succeq x^*_B x_{T\setminus B} g$ implies $y^*_A y_{T\setminus A} g \succeq y^*_B y_{T\setminus B} g$.

Savage's sure thing principle allows him, for any event E, to define a conditional preference relation $\succeq_E$ that is independent of the outcomes that would result if the state is not in E. That is, given P2, we can write $g \succeq_E f$ to mean $g_E h \succeq f_E h$ for some act h (and hence all acts) in $\mathcal{F}$. The decomposition property discussed in the introduction is immediately implied by P2 and, given P2, can be restated in terms of such 'independent' conditional preferences: if $g \succeq_A f$ and $g \succeq_{S\setminus A} f$ then $g \succeq f$. The implicit assumption that preferences are separable across events is, however, controversial.

Suppose, for example, that an individual is faced with the following two choice problems.

1. Choose between the act 1 (that is, the act that yields 1 in every state) and the act $5_A 0_B 1$ (the act that yields 5 if the state is in event A, 0 if the state is in event B and 1 otherwise).
2. Choose between the act $1_{A\cup B}0$ (the act that yields 1 if the state is in either event A or B, and zero otherwise) or $5_A 0$ (the act that yields 5 if the state is in event A, and zero otherwise).

These choices are subjective versions of those used to illustrate the Allais paradox or common consequence effect. In experiments where agents view $A \cup B$ as a fairly unlikely event and with A relatively much more likely than B, most individuals declare a strict preference for $1 \equiv 1_{A\cup B}1$ over $5_A 0_B 1$ in the first problem, but a strict preference for $5_A 0 \equiv 5_A 0_B 0$ over $1_{A\cup B}0$ in the second. This contradicts P2: conditional preference on the event $A \cup B$ depends on whether 1 or 0 is the consequence if this event did not occur. Many agents are able to 'rationalize' this pair of choices. For example, the agent might be wary of the subact $5_A 0_B$ when the outcome is 1 if neither A nor B occur and when the alternate subact $1_{A\cup B}$ is available. She anticipates that she will be especially unhappy should B occur knowing that she could have guaranteed herself 1 for sure. On the other hand, if the outcome is 0 if neither A nor B occur then choosing the subact $5_A 0_B$ does not introduce a new disappointing possible outcome. The possibility of the good outcome 5 now seems worth the risk. Explanations along these lines explicitly reject the consequentialist reasoning underpinning P2.

As discussed in the introduction, our approach in this paper is to replace P2 with a weaker axiom that still captures the intuition behind Savage's thought-experiment.

Weak Decomposability. For any pair of acts f and g in $\mathcal{F}$, and any event A in $\mathcal{E}$: $g_A f \succ f$ and $f_A g \succ f$ implies $g \succ f$.

In words, starting from the act f, if the agent is made better off by substituting g for f on A and she is also made better off by substituting g for f on $S\backslash A$, then she unconditionally prefers g to f. This axiom simply formalizes our interpretation of the example in the introduction, where f is the act 'not to buy the house', g is the act 'to buy', and A is the event of a Republican victory in the coming election.

Since we have dropped P2, the 'independent' conditional preference relation $\succeq_E$ will not in general be well-defined. That is, any preference relation 'conditional on the event E' will in general depend on the outcomes that would result if the state is not in E. For any event E and any act h, however, we can always construct a 'dependent' conditional preference relation $\succeq_{E,h}$ simply by setting $g \succeq_{E,h} f$ if $g_E h \succeq f_E h$. It is precisely because weak decomposability does not imply that conditional preferences are independent of what would happen if the event in question did not occur, that allows the weaker axiom to accommodate Allais choice patters like those described above. The first choice of act $1_{A\cup B}1$ over $5_A 0_B 1$ has implications for the 'dependent' conditional preference relation $\succeq_{A\cup B,1}$. It has no implication, however, for the 'dependent' conditional preference relation $\succeq_{A\cup B,0}$, and it is this preference relation that is relevant to the second Allais choice problem.

Although weak decomposability is written for a simple decomposition of the state space into two events, A and $S\backslash A$, the idea automatically extends to any finite partition.

Observation 1 *Suppose that preferences, $\succeq$, satisfy P1 and weak decomposability. Then, for all pairs of acts f and g, and any finite partition of the state space $\{E_1,\ldots,E_n\} \subset \mathcal{E}$, if $g_{E_i} f \succ f$ for all $i = 1,\ldots,n$ then $g \succ f$.*

Proof. We proceed by induction. Let $h^k := g_{E_1\cup\ldots\cup E_k} f$, so that $h^1 = g_{E_1} f$ and $h^n = g$. As an induction hypothesis, suppose that $h^k \succ f$. By assumption, this hypothesis holds for $k = 1$. For any $k \in \{1,\ldots,n-1\}$, we have $h^{k+1}_{E_{k+1}} f = g_{E_{k+1}} f \succ f$ (by assumption) and $f_{E_{k+1}} h^{k+1} = h^k \succ f$ (by induction hypothesis), hence $h^{k+1} \succ f$ (by weak decomposability). ■

A second attractive feature of weak decomposability is that it is automatically inherited by any conditional preference relation derived from $\succeq$. That is for any event E and any act h, the conditional preference relation $\succeq_{E,h}$ satisfies weak decomposability conditionally on the event E. Epstein & Le Breton (1993) and Sarin & Wakker (1997) argue that updated preferences should indeed still respect axioms imposed on parent preferences. This is a particularly natural property to require if we regard the parent preferences, $\succeq$, themselves to be derived from grandparent preferences, where updating has resulted from the resolution of some earlier (but unmodeled) uncertainty.

Observation 2 *If $\succeq$ satisfies weak decomposability then for any pair of disjoint events B and C in $\mathcal{E}$ and all acts f, g and h in $\mathcal{F}$: $g_B f_C h \succ f_{B\cup C}h$ and $f_B g_C h \succ f_{B\cup C}h$ implies $g_{B\cup C}h \succ f_{B\cup C}h$.*

Proof. Set $\widehat{f} := f_{B\cup C}h$, $\widehat{g} := g_{B\cup C}h$ and the result follows immediately from weak decomposability. ∎

Returning to our interpretation of Savage's thought experiment, the reader might object that we gave a special 'status-quo' role to the act 'not to buy the house'. In decomposing the comparison between 'to buy' and 'not to buy' into the events in which either the Republican or the Democrat won, we always used the act 'not to buy' on the event other than that under immediate consideration. For example we interpreted the statement the agent would prefer to buy the house if he knew the Republican would win to mean he prefers the act 'to buy if the Republican wins but not to buy otherwise' to the act 'not to buy regardless'. An equally reasonable interpretation is that he prefers the act 'to buy regardless' over the act 'not to buy if the Republican wins, but to buy otherwise'. The corresponding weak decomposability would then impose that, if in addition he prefers 'to buy regardless' over 'to buy if the Republican loses but not to buy otherwise' then the businessman should prefer the act 'to buy regardless' to the act 'not to buy regardless'. That is, starting from the act g, if the agent is made worse off by substituting f for g on A and she is also made worse off by substituting f for g on $S\backslash A$, then she unconditionally prefers g to f. The following proposition, however, shows that, given the Savage framework notions of "monotonicity" and "continuity", these two versions of weak decomposability are equivalent. Moreover, we can express the idea using weak or strict preference.

Proposition 1 *Suppose that Savage postulates P1, P3, and P6 hold. Then weak decomposability is equivalent to any of the following statements: for any pair of acts f and g in $\mathcal{F}$, and any event A in $\mathcal{E}$:*

1. *$g_A f \succeq f$ and $f_A g \succeq f$ imply $g \succeq f$;*
2. *$g_A f \succ f$ and $f_A g \succeq f$ imply $g \succ f$;*
3. *$g \succ f_A g$ and $g \succ g_A f$ imply $g \succ f$;*
4. *$g \succeq f_A g$ and $g \succeq g_A f$ imply $g \succeq f$;*

5. $g \succ f_A g$ and $g \succeq g_A f$ imply $g \succ f$.

Furthermore, weak decomposability implies:

6. $g_A f \sim f$ and $f_A g \sim f$ imply $g \sim f$,[5]

and this is equivalent to weak decomposability if in addition P6 holds.*

Consider two acts f and g, an event E and an outcome z. Suppose we can find outcomes x and y such that $y \succeq x$, $f_E x \sim z$ and $g_E y \sim z$. That is, to make the subact g on E indifferent to z we have to augment it with a (weakly) better outcome off E than we do for f on E. Then we can define a natural preference relation over the sub-acts f on E and g on E with respect to the outcome z, by the rule f on E is (weakly) preferred to g on E if $y \succeq x$. The following corollary says that, given weak decomposability, this induced preference relation satisfies P2. This is analogous to betweenness (in preferences over lotteries) implying independence within an indifference set.

Corollary 2 *Assume P1, P3, P6 and weak decomposability. Let A and B be disjoint, non-null events such that $(A \cup B)^c$ is also non-null. Suppose that $f_A h_B x \sim g_A h_B y \sim f_A h'_B x' \sim g_A h'_B y'$. Then $y \succeq x$ if and only if $y' \succeq x'$.*

Proof. Since $y \succeq x$, $f_A h_B x \sim g_A h_B y$ implies (by P3) $f_A h_B x \succeq g_A h_B x$. By hypothesis, $f_A h_B x \succeq f_A h'_B x'$. These two facts imply (by Proposition 1 part 1) $f_A h_B x \succeq g_A h'_B x'$. By hypothesis, $f_A h_B x \sim g_A h'_B y'$, hence $g_A h'_B y' \succeq g_A h'_B x'$. So, by P3, $y' \succeq x'$. ■

3 Representation

In this section, we provide a representation theorem for preferences that satisfy weak decomposability. The conditions we give for the sufficiency and necessity parts are not identical, so we break the result into two propositions.

In both results, all the conditions except P6 (and the corresponding property 5) can be met when the state space is finite. Indeed, the proofs require minimal modification to cover this case. We chose to retain P6 to keep our framework similar to that of Savage. For this section, we take $\mathcal{X}$ to be a compact interval in the real line, $[\underline{x}, \bar{x}]$.[6] This setting suggests a

[5] Property 6 can be regarded as an analogue to Chew & Epstein's (1989) "indifference separability" property, part of their axiomatization of betweenness preferences under risk.

[6] More generally, given P3, it is enough that the outcome space is rich enough for there to exist a utility function u defined over the set of constant acts (i.e. outcomes), and representing preferences over those acts, whose range is such an interval. For a discussion of related issues in the context of risk, see Grant, Kajii & Polak (1992).

natural notion of continuity. For a finite partition $\mathcal{P} = \{A_k : k = 1, \ldots, K\}$ of $\mathcal{S}$, denote by $\mathcal{F}^{\mathcal{P}}$ the set of acts that are measurable with respect to $\mathcal{P}$. This set is naturally identified with $\mathcal{X}^K$, a subset of $\mathbb{R}^K$. We shall refer to a finite partition $\mathcal{P}$ as non-null with respect to the preference relation $\succeq$, if every element A in $\mathcal{P}$ is not null. The preference relation $\succeq$ is said to be continuous with respect to a non-null partition $\mathcal{P}$ if it induces a continuous relation on $\mathcal{F}^{\mathcal{P}}$.

The following proposition gives conditions for a representation V that are sufficient for the induced preferences to satisfy weak decomposability.

Proposition 3 *Let $\mathcal{X} = [\underline{x}, \bar{x}]$. For a state space $\mathcal{S}$, a set of events $\mathcal{E}$, and a function $\varphi : \mathcal{X} \times \mathcal{E} \times \mathcal{X} \rightarrow \mathbb{R}$, let $\mathcal{N} := \{A \in \mathcal{E} : \varphi(x, A, w) = 0 \text{ for all } x, w \text{ in } \mathcal{X}\}$. Suppose that $\mathcal{S}$, $\mathcal{E}$, and φ satisfy the following properties:*

1. *φ is continuous in its first and third arguments;*
2. *for all events A in $\mathcal{E} \backslash \mathcal{N}$, $\varphi(., A, .)$ is increasing in the first argument and is decreasing in the third argument;*
3. *for all events A in $\mathcal{E}$, and all x in $\mathcal{X}$, $\varphi(x, A, x) = 0$;*
4. *φ is state additive; that is, for all pairs of events A and B in $\mathcal{E}$, and all x and w in $\mathcal{X}$, if $A \cap B = \emptyset$ then $\varphi(x, A, w) + \varphi(x, B, w) = \varphi(x, A \cup B, w)$;*
5. *φ is small-event continuous; that is, for all outcomes x and w in $\mathcal{X}$ with $x \neq w$, and all $\varepsilon > 0$, there exists a finite partition $\mathcal{P} = \{A_k : k = 1, \ldots, K\}$ of $\mathcal{S}$, $\mathcal{P} \subset \mathcal{E}$, such that $|\varphi(x, A_k, w)| < \varepsilon$ for all $A_k \in \mathcal{P}$.*

And, for all simple acts f in $\mathcal{F}$, let $V(f)$ be the (unique) implicit solution to:

$$\sum_{x \in \mathcal{X}} \varphi\left(x, f^{-1}(x), V(f)\right) = 0 \tag{1}$$

Then the relation $\succeq$ induced by the functional V satisfies P1, P3, P6, continuity with respect to all non-null partitions,[7] and weak decomposability.

To help understand this representation, first notice that we can write the standard subjective expected utility model in this form. For example, set $S := [0, 1]$ and suppose that

$$\bar{V}(f) = \int_s u(f(s))\, \mu(ds) = \sum_{x \in \mathcal{X}} u(x)\, \mu \circ f^{-1}(x), \tag{2}$$

[7] That is, non-null with respect to $\succeq$.

where u is a continuous increasing utility index and where μ is a strongly continuous finitely additive probability measure.[8] We can rewrite expression (2) as

$$\sum_{x\in\mathcal{X}} \left(u\left(x\right) - \bar{V}(f)\right) \mu \circ f^{-1}\left(x\right) = 0.$$

If we now set $\bar{\varphi}(x, E, v) := (u(x) - v)\mu(A)$, then this reduces to expression (1). It is easy to check that $\bar{\varphi}$ satisfies the six properties of Proposition 3. In this special case, however, the preferences induced by $\bar{V}$ also satisfy P2 and P4. We want to allow more general preferences.

For a more general special case, suppose that $V^*(f)$ is given by the solution, v, to

$$\int_{s\in S} \psi\left(f(s), s, v\right) ds = 0, \tag{3}$$

where the function $\psi : \mathcal{X} \times s \times \mathcal{X} \to \mathbb{R}$ satisfies three properties analogous to properties 1, 2 and 3 above: ψ is continuous in its first and third arguments, increasing in the first and decreasing in the third, and $\psi\left(x, s, x\right) = 0$ for all s and x. The function V^* resembles an implicit linear representation of a betweenness preference relation over lotteries (see, for example, Chew (1989)). For each v and s, we can think of $\psi\left(., s, v\right)$ as assigning a 'utility' to each outcome, where that utility depends on the state s in which the outcome occurs. Then, for fixed v, the left side of (3) is the expected 'state-dependent utility' of the act f with respect to Lesbesgue measure.

Just like their betweenness representation analogues, the three properties of ψ ensure that the solution v exists and is unique, thus the preferences induced by V^* satisfy P1 (order). The third property ensures that, for all constant acts x, $V^*(x) = x$. Monotonicity in the first argument thus ensures that the induced preferences satisfy P3 (eventwise monotonicity). Small event continuity (P6) and continuity with respect to non-null partitions are both immediate in this case, so it only remains to show weak decomposability. Monotonicity in the third argument ensures, as with betweenness representations, $V^*(g) \geq V^*(f)$ if and only if $\int_{s\in S} \psi\left(g(s), s, V^*(f)\right) ds \geq 0$. Thus, $V^*(g_A f) \geq V^*(f)$ implies $\int_{s\in A} \psi\left(g(s), s, V^*(f)\right) ds \geq 0$, and $V^*(f_A g) \geq V^*(f)$ implies $\int_{s\in S\setminus A} \psi\left(g(s), s, V^*(f)\right) ds \geq 0$. And, combining these two implications, we are done. This is essentially the idea of the proof of the proposition.

To see that this example is indeed a special case of the functional form in the proposition, rewrite (3) as

$$\sum_{x\in\mathcal{X}} \int_{s\in f^{-1}(x)} \psi\left(x, s, V^*(f)\right) ds = 0.$$

[8] A measure with this property is sometimes referred to as an atomless measure, but this is confusing since we are concerned with finitely additive measures: in probability theory, an atomless measure refers to a measure μ for which there is no event E with $\mu\left(E\right) > 0$ such that $E' \subset E$ implies either $\mu\left(E'\right) = 0$ or $= \mu\left(E\right)$. An atomless measure is strongly continuous if it is countably additive, but not necessarily so if it is only finitely additive. See Bhaskara Rao & Bhaskara Rao (1983).

If we now set $\varphi^*(x,E,v) := \int_{s\in E} \psi(x,s,v)\,ds$, then this reduces to expression (1). The function φ^* inherits properties 1, 2 and 3 from ψ, and satisfies state-additivity (property 4) and the analogue of small-event continuity (property 5) by construction.

The functional form in Proposition 3 is more general than V^*.[9] In particular, state additivity does not imply the existence of a measure (integral) representation like those in expressions (2) and (3) above. Wakker & Zank (1996) consider preferences which satisfy all of Savage's first six axioms except P4. Such preferences need not admit a separation between the 'probability' and the 'utility' of outcomes. They show that, in general, in the absence of an identifiable probability measure, no integration operation can be defined. Since class of preferences in Proposition 3 are even less restricted than those of Wakker & Zank, we must make do with the additive form of expression (1).[10]

The following proposition gives conditions for a representation V that are necessary if the represented preferences satisfy weak decomposability and the same technical conditions as above.[11]

Proposition 4 *Let $\mathcal{X} = [\underline{x}, \bar{x}]$. Assume that $\succeq$ satisfies P1, P3, P6 and weak decomposability. Suppose further that $\succeq$ satisfies continuity with respect to any non-null partition $\mathcal{P} = \{A_k : k = 1, \ldots, K\} \subset \mathcal{E}$ with $K > 3$. Then there exists a function $\varphi : \mathcal{X} \times \mathcal{E} \times (\underline{x}, \bar{x}) \to \mathbb{R}$ with the properties: if A is null then $\varphi(x, A, w) = 0$ for all x, w in $\mathcal{X}$; and*

1. *φ is continuous in its first arguments;*
2. *for all events A in $\mathcal{E}\backslash\mathcal{N}$, and all w in $(\underline{x}, \bar{x})$, $\varphi(., A, w)$ is increasing in the first argument*
3. *for all events A in $\mathcal{E}$, and all x in $\mathcal{X}$, $\varphi(x, A, x) = 0$;*
4. *φ is state additive;*
5. *φ is small-event continuous with respect to all outcomes x and w in $(\underline{x}, \bar{x})$, $x \neq w$.*

[9] For another example of a function that satisfies properties (1)-(6) consider $\tilde{\varphi}(x, A, v) := \int_{s\in A} [x - v]\, p(ds, v)$.

[10] Even when an integral representation exists, we should not identify the measure over states as the subjective beliefs of the agent. Consider, for example, the functional form (3) above. Let $\mu(\cdot)$ be a probability measure with density μ'. Set $\hat{\psi}(x,s,v) := \psi(x,s,v)\,\frac{1}{\mu'(s)}$. Then $\int_{s\in S} \hat{\psi}(f(s), s, v)\, d\mu(s) = 0$ induces the same preferences as expression (3). Since there is no particular reason for choosing ψ over $\hat{\psi}$, this shows that any such μ could serve as beliefs.

[11] The main difference between the two propositions is that, while it is sufficient for the functions φ to be everywhere decreasing in the third argument, we do not know if it is necessary.

such that the relation $\succeq$ on simple acts (excluding the two constant acts $\underline{x}$ and $\bar{x}$) is represented by a utility function V given by the rule $\sum_{x\in\{y:f^{-1}(y)\notin\mathcal{N}\}}\varphi\left(x,f^{-1}(x),V(f)\right)=0$.[12]

The idea of the proof is, loosely, as follows. Recall from our discussion of Corollary 2 that we can think of how much we have to 'augment' each subact to bring it into an indifference set, as inducing a natural ordering over subacts with respect to that indifference set. Moreover, this ordering respects P2. Thus, using Segal's (1992, Theorem 2) result, the ordering has an additively separable representation. We then show that if two subacts agree on the intersection of two events, then the representation (still with respect to a given indifference set) agrees on that intersection up to affine transformations.. From there, the proof adapts methods of Chew & Epstein (1989) and Wakker & Zank (1996). Notice, however, that unlike the former we do not assume probabilistic sophistication (or even P4), and unlike the latter we do not assume P2.

4 Computation, Planning and Decentralization

In this section, we show that weak decomposability is essentially equivalent to three other properties. Each of these properties is useful in a different context. The first concerns the computing of optimal plans; the second concerns whether parts of optimal plans are interchangeable; and the third concerns decentralizing choice to the 'agents' in agent-normal form games. This exercise is very much in the spirit of Gul & Lantto (1990), and indeed the second property is a translation of theirs.

Computing Optimal Plans. In many areas of economics, we place conditions on problems to make it easier to compute a solution. For example, we often assume some kind of convexity condition in maximization problems to ensure that local necessary conditions for optimality are in fact globally sufficient. Weak decomposability can also be thought of as a condition under which it is enough to check 'local' necessity conditions.

Suppose there are a finite number of possible actions, and that plans assign an action to each of some (mutually exclusive and exhaustive) finite set of events. For example, the actions to be chosen might be four available choices of drink: orange, red wine, white wine and beer. The events which the agent does not control might be three possible main courses: meat, fish or vegetarian. A drink plan assigns a drink to each of three possible main courses; say, orange juice if meat, red wine if fish, and beer if vegetarian. A natural computational

[12] Notice that the exclusion of $\underline{x}$ and $\bar{x}$ is without loss of generality since by P3 (eventwise monotoncity) these two acts cannot be certainty equivalents for any other acts.

algorithm is to pick some candidate plan at random, then check if it can be improved at one event, trying each event and each alternative at that event in turn. If such a 'local' improvement is found, adopt the adjusted plan as the new candidate and repeat. Once the a candidate plan can not be improved on any single event, the algorithm stops and this plan is chosen. In our example, perhaps, first orange with meat would be switched to red wine. Then, perhaps, the red wine with fish would be switched to orange, and finally this orange switched to white wine.

To formalize this idea, the following notation will be useful. For any finite set of simple acts $\mathcal{H} \subset \mathcal{F}$ and any finite partition $\mathcal{P} := \{A_1, \ldots, A_N\}$ of the state space, let $H(\mathcal{H}, \mathcal{P})$ be the set of all acts of the form $h(1)_{A_1}h(2)_{A_2} \ldots h(N-1)_{A_{N-1}}h(N)$ with $h(i)$ in $\mathcal{H}$ for all $i = 1, \ldots, N$. In the food and drink example, $\mathcal{P}$ could be the categories of main course (meat, fish, etc.), and $\mathcal{H}$ could be the set of acts corresponding to uncontingent drink allocations (such as "beer with everything"). In this case, the set of acts generated by all possible food and drink combinations is equal to $H(\mathcal{H}, \mathcal{P})$.

Given some finite set F of feasible acts, a finite partition of the event space $\mathcal{P} := \{A_1, \ldots, A_N\}$, and a subset $\mathcal{H} := \{h^1, \ldots, h^M\}$ of F such that $F = H(\mathcal{H}, \mathcal{P})$,[13] consider the following algorithm.

Algorithm A

(0) *Set* $g = h^1$.

(1) *Set* $i = 0$.

(2) *Set* $i = i + 1$ *and set* $j = 0$.

(3) *Set* $j = j + 1$.

(4) *If* $h^j_{A_i}g \succ g$ *then set* $g = h^j_{A_i}g$ *and go to* (1).

(5) *If* $j < M$ *then go to* (3).

(6) *If* $i < N$ *then go to* (2).

(7) *Set* $f^* = g$ *and end.*

Given P1, since the problem is finite, we know that this algorithm will stop. We want to ensure that where-ever it stops is an optimum. Returning to our food and drink example, a

[13] If F were a strict subset of $H(\mathcal{H}, \mathcal{P})$ then the algorithm below could stop at an act that is not feasible. The restriction $F = H(\mathcal{H}, \mathcal{P})$ is not, however, as stringent as it might appear. In our example, if 'beer with meat' is impossible then $\mathcal{H}$ may no longer include the act corresponding to 'beer with everything'. But, if we replace it by the act corresponding to 'beer with everything except meat, then red wine", then equality between F and $H(\mathcal{H}, \mathcal{P})$ is restored.

necessary condition for a particular drink plan g to be optimal is that there is no plan that differs from g on only one main course and which is better. This is what the algorithm checks: $f^* \succeq h_A f^*$ for all h in $\mathcal{H}$ and all A in $\mathcal{P}$. In a sense this is a local condition: it considers deviations from the candidate optimal act at only one 'place' at a time. The property below ensures that this local necessary condition is always globally sufficient.

A-Sufficiency *The preference relation $\succeq$ satisfies* A-sufficiency *if, for all finite sets of feasible acts $F \subset \mathcal{F}$, all finite partition of the event space $\mathcal{P} \subset \mathcal{E}$ and all subsets $\mathcal{H}$ of the feasible acts such that $F = H(\mathcal{H}, \mathcal{P})$, the output f^* of Algorithm A is optimal.*

To see why A-sufficiency is computationally convenient consider testing whether a given candidate act is optimal. A crude sufficiency check would involve checking the candidate against all $\#F - 1$ other feasible acts. In our example, with four drinks and three main courses, there are $4^3 - 1$ ($\#\mathcal{H}^{\#\mathcal{P}} - 1$) such candidates. Given A-sufficiency, however, we need only check the local necessary condition for optimality. This involves only 4×3 preference comparisons (or 3×3 if we never check an act against itself). More generally, it involves only $(\#\mathcal{H} - 1) \times \#\mathcal{P}$ comparisons.

Interchangeable Optimal Plans. Gul & Lantto (1990) consider normative rules for an agent's choices within and between dynamic decision trees. Their aim, following Machina (1989), is to weaken the standard consequentialist assumption, while still retaining some degree of consistency across choices within trees. Among the normative restrictions they suggest is a property they call dynamic programming solvability (DPS). The motivation for DPS is similar to the computational considerations discussed above. To illustrate the idea, they give the example of an agent who has to decide how to go to work. The options are to walk, to drive, to bike, or to take the bus. Suppose the following two plans are optimal: (1) drive if it rains, bike if it is sunny; and (2) take the bus if it rains, walk if it is sunny. Then, they argue, the following plans of actions should also be optimal: (3) drive if it rains, walk if it is sunny; and (4) take the bus if it rains, bike if it is sunny. That is, dynamic programming solvability implies that "shuffling" two optimal plans of actions produces another optimal plan of action.

This seems a desirable property for preferences to have in dynamic decision problems. Unlike consequentialism, it does not imply that the agent's choice at each final decision node is independent of what her choices would have been at other, unrealized, nodes. But it does mean that, if there is more than one optimal overall plan, her choice at each final decision node does not depend on which optimal plan she would have followed at the other nodes. In this sense, DPS is a decomposition property, albeit in another context, and it would be nice if it were related to weak decomposability. To explore any relation, however, we have first to translate their axiom on dynamic choices under objective risks to an axiom on static preferences under subjective uncertainty. We trust that the following property captures at least some of Gul & Lantto's original intuition. Let $H(\mathcal{H},\mathcal{P})$ be defined as above.

DPS* *The preference relation* $\succeq$ *satisfies* DPS* *if, for any finite partition* $\mathcal{P} := \{A_1, \ldots, A_n\}$ *of the state space* $\mathcal{S}$*, and for any pairs of acts* f *and* g *in* $\mathcal{F}$*: if* $f \sim g$ *and* $f \succeq h$ *for all acts* h *in* $H(\{f,g\},\mathcal{P})$*, then* $f \sim h$ *for all acts* h *in* $H(\{f,g\},\mathcal{P})$*.*

Decentralization in Agent-Normal Forms. Consider a two-player game in which player one moves first; player two moves second; then the game ends. There might also be some background uncertainty (that is, a move by nature), but let us assume that nature's moves are not observed by either player prior to their choices being made. In the following, we will only be concerned with player two's point of view. Borrowing notation from above, let $\mathcal{S}$ be the set of states reflecting player two's uncertainty both about nature and about player one's move. Let outcomes (which depend on nature and on both actions) be elements of $\mathcal{X}$. Suppose that player two has (finite) N information sets, and (finite) M actions available at each set.[14] The game need not be of perfect information; many actions of player one could be associated with each information set.

Consider the normal form of the game. Each (pure) strategy for player two selects an action at each of her information sets, contingent on that set being reached. The strategy thus induces an act f that assigns an outcome to each state. To stay within the framework of this paper, assume that all such induced acts have finite outcome sets. Let F be the set of such acts that can be induced by some strategy. If player two's preferences over $\mathcal{F}$ (the set of all finite outcome acts) are given by $\succeq$, then it is natural to define a strategy as optimal for player two if it induces an act f^* such that $f^* \succeq f'$ for all f' in F. In this case we refer to f^* as an optimal in F.

Now consider the agent-normal form of this game. Player two has N agents (indexed by i) each of whom has M possible actions (indexed by j). Let $\mathcal{P} :=\{A_1,\ldots,A_N\} \subset \mathcal{E}$ be the partition of $\mathcal{S}$ corresponding to Player two's information sets. We can think of each agent i choosing player two's action for her ith information set, and hence choosing the 'subact' on the event A_i. We will assume throughout that each agent i of player two inherits the same preferences $\succeq$ over the acts in $\mathcal{F}$. So, in particular, all player two's agents agree about (the

[14] Since we can always duplicate actions, it is without loss of generality to assume the same number at each set.

'undivided') player two's optimal strategies.[15]

In the agent normal form, implicitly, the choices of the agents of player two are made simultaneously to each other and to player one's choice. In such games, under the standard assumptions (in particular, with Savage's P2), agent i of player two need not think about the choices of other agents of player two when making her choice: the problem is separable. Without P2, however, the best choice of agent i may depend on what agent i' chooses. In general, even though agents i and i' agree about the optimum overall strategy, when they choose their actions separately, they may fail to coordinate on such a strategy. Each agent of player two could be choosing the best available action (and hence sub-act) given the choices of the other agents of player two. But, this Nash behavior need not add up to an optimal strategy for player two overall. For the agent normal form to an appropriate analytical tool, we would like to rule out such coordination failures.

To formalize this idea, let h^j denote the act in $\mathcal{F}$ induced by player two choosing the jth action at every information set. Let $\mathcal{H} := \{h^1, \ldots, h^M\}$. Then the set of feasible acts, F, is the set of acts of the form $f := h(1)_{A_1}h(2)_{A_2}\ldots h(N-1)_{A_{N-1}}h(N)$ with $h(i)$ in $\mathcal{H}$ for all $i = 1, \ldots, N$. By construction (using our earlier notation), we have $F = H(\mathcal{H}, \mathcal{P})$. It is then natural to define a strategy of player two as 'Nash among her agents', if it induces an act $\hat{f}$ such that if $\hat{f} \succeq h^j_{A_i}\hat{f}$ for all agents i, and actions j. In this case, we refer to the act $\hat{f}$ as Nash in $H(\mathcal{H}, \mathcal{P})$.

Agent-Normal Form Decentralizability. *A preference relation $\succeq$ satisfies* agent-normal form decentralizabilty *if, for all finite partitions of the state space, $\mathcal{P} \subset \mathcal{E}$, and all finite sets of acts $\mathcal{H} \subset \mathcal{F}$, if f is Nash in $H(\mathcal{H}, \mathcal{P})$ then f is optimal in $H(\mathcal{H}, \mathcal{P})$.*

Equivalence. The following proposition states formally the equivalence of weak decomposability and the three properties above. The result is not technically difficult or surprising, but it does suggest that weak decomposability is an important normative property. Notice that the equivalences below do not require the agent to be probabilistically sophisticated. They do not even require Savage's P4 axiom. Also, although we assume both continuity axioms, P6* is only used to show that DPS* implies weak decomposability.

Proposition 5 *Suppose that the preference relation $\succeq$ satisfies* P1, P3, *and* P6 *and* P6*. *Then the following are equivalent properties of $\succeq$:*

---

[15] This assumption requires more than just that each agent have the same preferences over outcomes (or even over lotteries on outcomes). It is as if we require each agent to have the same (not necessarily probabilistic) conjecture about player one's actions. However, since these agents are all manifestations of the same player, this restriction seems reasonable.

(1) *Weak Decomposability;*
(2) *A-sufficiency;*
(3) *DPS*;*
(4) *Agent-normal form decentralizability.*

For intuition, recall that Savage's sure thing principle rules out 'local' optima that are not 'global' by imposing separability across events. Weak decomposability relaxes separability but achieves the same end by limiting the way in which preferences conditional on each event can depend on what would happen if some other event were to occur. Loosely speaking, problems arise when the subact f on A is a 'complement' of the subact g on $S\backslash A$. Consider Diamond's problem of a mother allocating an indivisible gift to one of her two children, Fred and Gail. For fairness reasons, the subact 'give the gift to Fred if Australia wins the test match' may be a complement to the subact 'give the gift to Gail otherwise'. If so, in the obvious notation, we could have $f_A g \succ f$ and $f_A g \succ g$, so that $f_A g$ is a local optimum, but still have $g_A f \succ f_A g$.[16] Weak decomposability rules out this degree of 'complementarity' across events.[17]

5 Probabilistic Sophistication and Betweenness

In this section, we first show that weak decomposability (in conjuction with Savage's other axioms except P2) does not imply probabilistic sophistication. We then show that, if probabilistic sophistication is assumed, then weak decomposability is equivalent to betweenness.

Probabilistic Sophistication. Savage's original sure thing principle was part of an axiomatization not only of separable preferences over lotteries (expected utility) but also of additive subjective beliefs (probabilistic sophistication). Recent work by Machina & Schmeidler (1992) and others[18] has shown that additive subjective beliefs can be axiomatized without requiring the agent's preferences over lotteries to obey the expected utility hypothesis. Formally they define probabilistic sophistication as:

Definition *A preference relation is said to be* probabilistically sophisticated, *if there exists*

[16] The use of a random mechanism to allocate an indivisible good in an equitable or fair way is an old idea. For example it appears in Hobbes (1651, Chapter XV, p.165). We thank Mamoru Kaneko for bringing this reference to our attention.

[17] Johnsen & Donaldson (1985) consider perverse preferences rather like these in a dynamic context. They use what they call "conditional weak independence" to rule them out. Peter Klibanoff suggested the term complementarity.

[18] See Epstein & LeBreton (1993) and Grant (1995).

a finitely additive probability measure μ *on* $\mathcal{E}$ *such that for any pair of acts* f *and* g*, if* $\mu \circ f^{-1}(x) = \mu \circ g^{-1}(x)$ *for all* x *in* $f(\mathcal{S}) \cup g(\mathcal{S})$ *then* $f \sim g$*.*

Probabilistic sophistication means that we can represent the individual's beliefs over the states of the world with a probability measure and, moreover, we can separate those beliefs from her 'risk preferences'. To see this, notice that we can use the measure μ that represents her beliefs, to map acts into $\mathcal{L}_0$, the set of lotteries with finite support, as follows:

$$f \mapsto P, \text{ where } P(x) = \mu \circ f^{-1}(x) \text{ for all } x \text{ in } f(\mathcal{S})$$

We can then identify the individual's 'risk preferences' with the induced relation over lotteries with finite support; that is, we can define an induced relation $\succeq_{\mathcal{L}_0}$ over lotteries with a finite support by the rule that for any two lotteries P and Q, $P \succeq_{\mathcal{L}_0} Q$ implies there exists two acts f and g such that $f \succeq g$, $\mu \circ f^{-1}(x) = P(x)$ and $\mu \circ g^{-1}(x) = Q(x)$ for all x in $f(\mathcal{S}) \cup g(\mathcal{S})$. Probabilistic sophistication ensures that the induced relation is transitive. Moreover knowledge of $\succeq_{\mathcal{L}_0}$ and μ enables the analyst to recover all of $\succeq$ since for any pair of acts f and g which are mapped by μ to P and Q respectively, we may correctly infer that $f \succeq g$ if and only if $P \succeq_{\mathcal{L}_0} Q$.

None of the discussion in the previous sections depended on whether or not the agent was probabilistically sophisticated, but there remains the issue as to whether substituting weak decomposability for P2 still implies probabilistically sophistication. We show by counter-example that it does not. That is, an agent's preferences can be decomposable without necessarily being based on a Bayesian system of beliefs.

Example 1 *Let the state space* $\mathcal{S}$ *be the interval* $[0,1]$ *with Lebesgue measure, and let* $\{R, W, B\}$ *be a partition of* $\mathcal{S}$ *with* $R = [0, 1/3)$ *and* $B = [2/3, 1]$*. The set of outcomes* $\mathcal{X}$ *is taken to be* $[0,1]$*. The preference relation* $\succeq$ *is represented by a function* $V : \mathcal{L}_0 \to [0,1]$*, where* V *is implicitly defined as* $\int_0^1 \varphi(f(s), s, V(f))\, ds = 0$*, with*

$$\varphi(x, s, v) = \begin{cases} x - v & \text{if } (x > v \text{ and } s \in R) \text{ or } (x \leq v \text{ and } s \notin R) \\ (1-a)(x-v) & \text{if } (x \leq v \text{ and } s \in R) \text{ or } (x > v \text{ and } s \notin R) \end{cases}$$

where $a \in (0,1)$*.*

These preferences can be thought of as a state-dependent version of Gul's (1991) disappointment aversion preferences in which, on the event R, 'good' outcomes are relatively overweighted but, off R, 'bad' outcomes are relatively overweighted.[19]

[19] Alternatively, for two-outcome acts, it can be shown that the preferences have a Choquet integral

Observation 3 *An agent with the preference relation defined in Example 1 satisfies weak decomposability and each of Savage's first six postulates except P2. Yet, she is not probabilistically sophisticated. In fact, these preferences can accommodate the Ellsberg paradox.*

Proof. Since these preferences conform with the form in Proposition 3, they satisfy weak decomposability, P3 and P6. In the appendix we show that they also satisfy P4. To see that these preferences both violate P2 and are not probabilistically sophisticated, recall the choices in Ellsberg's (1961) proposed experiment. An agent must bet on the draw of a ball from an urn containing red, white and blue balls. She only knows that a third of the balls are red. Let R (respectively, W, B) be the event that the color of the drawn ball is red (resp. white, blue). Fixing two outcomes x and y, $x > y$, the acts considered are

Act	R	W	B
$f' := x_R y$	x	y	y
$g' := x_W y$	y	x	y

Act	R	W	B
$f'' := x_{R\cup B} y$	x	y	x
$g'' := x_{W\cup B} y$	y	x	x

Ellsberg predicted that the typical agent would prefer f' to g', but prefer g'' to f'', thus exhibiting 'uncertainty aversion'. Such preferences violate P2 since $f'_{R\cup W} y \succ g'_{R\cup W} y$ but $g'_{R\cup W} x \succ f'_{R\cup W} x$. They also violate probabilistic sophistication since $f' \succ g'$ implies $\Pr(R) > \Pr(W)$, while $g'' \succ f''$ implies $\Pr(R) + \Pr(B) < \Pr(W) + \Pr(B)$.

The preferences given in Example 1, however, are consistent with Ellsberg's prediction. To see this, let $v' := V(x_R y) = (x + 2y)/3$ and $v'' := V(x_{W\cup B} y) = (2x + y)/3$. Then,

$$\begin{aligned} & \int_0^1 \varphi\left(g'(s), s, v'\right) ds - \int_0^1 \varphi\left(f'(s), s, v'\right) ds \\ = \; & \frac{1}{3}(1-a)\left(y - v'\right) + \frac{1}{3}(1-a)\left(x - v'\right) + \frac{1}{3}\left(y - v'\right) \\ & -\frac{1}{3}\left(x - v'\right) - \frac{1}{3}\left(y - v'\right) - \frac{1}{3}\left(y - v'\right) \\ = \; & -\frac{a}{3}\left(y + x - 2v'\right) = -\frac{a}{9}(x - y) < 0. \end{aligned}$$

Similarly

$$\begin{aligned} & \int_0^1 \varphi\left(g''(s), s, v''\right) ds - \int_0^1 \varphi\left(f''(s), s, v''\right) ds \\ = \; & -\frac{a}{3}\left(y + x - 2v''\right) = \frac{a}{9}(x - y) > 0. \end{aligned}$$

representation where the non-additive measure, ν on S, is given by

$$\nu(A) = \frac{[\mu(A \cap R) + \mu(A \cap (S \backslash R))(1-a)]}{[1 - a(\mu(A \cap (S \backslash R)) + \mu((S \backslash A) \cap R))]}.$$

Hence, $V(x_R y) > V(x_W y)$ and $V(x_{R\cup B} y) < V(x_{W\cup B} y)$ as desired. ■

Betweenness. Although weak decomposability does not itself imply that the agent has probabilistic beliefs, if we assume probabilistic sophistication then weak decomposability is equivalent to the Chew-Dekel betweenness property. That is, if for any two lotteries P and Q, $P \succ_{\mathcal{L}} Q$, then any probability mixture (i.e., convex combination) of the two lotteries, $\lambda P + (1-\lambda) Q$, we have $P \succ_{\mathcal{L}} \lambda P + (1-\lambda) Q$ and $\lambda P + (1-\lambda) Q \succ_{\mathcal{L}} Q$.

Proposition 6 *Suppose the preference relation $\succeq$ over $\mathcal{F}$ is probabilistically sophisticated, and satisfies P1, P3, and P6. Let μ be the associated finitely additive probability measure on $\mathcal{E}$. Then the following statements are equivalent:*

(i) $\succeq$ satisfies weak decomposability.

(ii) $\succeq_{\mathcal{L}_0}$ is represented by a continuous utility function V which is both quasi-convex and quasi-concave in probability mixtures (that is, $\succeq_{\mathcal{L}_0}$ satisfies betweenness).

For intuition, recall that Gul & Lantto's (1990) original DPS property is equivalent to betweenness, and weak decomposability is equivalent to at least a version of DPS. Alternatively, just as betweenness implies independence within an indifference set, by Corollary 2, weak decomposability induces a preference relation with respect to an indifference set that satisfies P2. Alternatively again, recall that weak decomposability rules out 'complementarity' across events. The linear indifference sets of betweenness preferences, similarly rule out such complementarity.

Combining the earlier work of Chew (1983, 1989), Dekel (1986) Machina & Schmeidler (1992) and Epstein & Le Breton (1993), with Proposition 6 yields the immediate corollary that weak decomposability forms part of an axiomatization of betweenness theory.

Corollary 7 *Suppose that the set of outcomes $\mathcal{X}$ is the interval $[0,1]$ and that $x > y$ implies $x \succ_{\mathcal{X}} y$. Then the following statements are equivalent.*

1. *The preference relation $\succeq$ over $\mathcal{F}$ satisfies P1, P3, P6, weak decomposability and $P4^C$.*

2. *There exists a finitely additive, strongly continuous probability measure μ on $\mathcal{E}$ and a function $V : \mathcal{F} \to \mathbb{R}$ that represents $\succeq$, such that V is implicitly defined by $\int_{\mathcal{X}} v(x, V(f)) \mu \circ f^{-1}(dx) = 0$ where $v : \mathcal{X} \times \mathbb{R} \to \mathbb{R}$ is increasing in its first argument.*

Appendix

Proof of Proposition 1. Let b denote a best outcome and w denote a worst outcome in the range of acts in question. First, observe that P3 and P6 imply the following continuity property. If $g \succ f$ and A is a non-null event with $g(s) \succ_{\mathcal{X}} w$ for all s in A, then there exists a non-null event $E \subset A$ such that $g \succ w_E g \succ f$. To see this, notice that, by P6, there exists a finite partition $\{E_1, \ldots, E_n\}$ of $\mathcal{S}$, such that $w_{E_i} g \succ f$ for all $i = 1, \ldots, n$. As the partition is finite and A is non-null, there exists an element, say E_1, that has a non-null intersection with A. Set $E = A \cap E_1$. Since, by construction, g is 'better' than w on A, by P3 (exploiting the fact that g has finite range), $g \succ w_E g \succeq w_{E_1} g \succ f$ as desired. By a similar argument, if $g \succ f$ and A' is a non-null event with $b \succ_{\mathcal{X}} f(s)$ for all s in A', there exists a non-null event $E' \subset A'$ such that $g \succ b_{E'} f \succ f$.

Weak decomposability implies (1): Suppose $g_A f \succeq f$ and $f_A g \succeq f$ but, contrahypothesis, $f \succ g$. If A is null then $g \sim f_A g$, which contradicts $f_A g \succeq f$ and $f \succ g$. Similarly, if $\mathcal{S} \backslash A$ is null then $g \sim g_A f$ which contradicts $g_A f \succeq f$ and $f \succ g$. So assume both A and $\mathcal{S} \backslash A$ are not null. Let $B' := \{s \in A : b \succ_{\mathcal{X}} g(s)\}$. If B' is null then, by P3, $g \succeq f_A g$, which again contradicts $f_A g \succeq f$ and $f \succ g$. So assume B' is not null. Since $f \succ g$, applying the observation above (with B' in the role of A'), there exists an event $E' \subset B'$ such that $f \succ b_{E'} g$. Similarly, if the set $B'' := \{s \in \mathcal{S} \backslash A : b \succ_{\mathcal{X}} g(s)\}$ is null then, by P3, $g \succeq g_A f$, which contradicts $g_A f \succeq f$ and $f \succ g$. So assume that B'' is also not null. By construction, $b \succ_{\mathcal{X}} [b_{E'} g](s)$ for all s in the non-null set B'' and $f \succ b_{E'} g$. Applying the above observation again (with B'' now in the role of A'), there exists an event $E'' \subset B''$ such that $f \succ b_{E'} b_{E''} g$. Now, set $g' := b_{E'} b_{E''} g$. By P3, we have $g'_A f = b_{E'} g_{A \backslash E'} f \succ g_A f$, and $f_A g' = f_A b_{E''} g \succ f_A g$. So we get $g'_A f \succ f$ and $f_A g' \succ f$ but $f \succ g'$, a contradiction to weak decomposability. □

(1) implies (2): Suppose $g_A f \succ f$ and $f_A g \succeq f$ but, contrahypothesis, $f \succeq g$. Since $g_A f \succ f$, the event A is not null. If $g(s) \sim_{\mathcal{X}} w$ on A except for a null event, then, by P3, $g_A f \succ f$ could not hold. So assume that $g \succ_{\mathcal{X}} w$ on a non-null event in A. Applying the observation, find a non-null event $E \subset A$ with $w_E g_{A \backslash E} f \succ f$. Since $f \succeq g$, by P3, $f \succ w_E g$. But if we set $g' = w_E g$, we get a contradiction to (1). □

(2) implies (3): Suppose $g \succ f_A g$ and $g \succ g_A f$ but, contrahypothesis, $f \succeq g$. If $f_A g \succeq g_A f$, we get $f \succeq g \succ f_A g \succeq g_A f$. Set $\hat{f} := f_A g$ and $\hat{g} := g_A f$, thus $\hat{f}_A \hat{g} = f$ and $\hat{g}_A \hat{f} = g$. So $\hat{f}_A \hat{g} \succeq \hat{g}_A \hat{f} \succ \hat{f} \succeq \hat{g}$, a contradiction to (2). Conversely, if $g_A f \succ f_A g$, we get $f \succeq g \succ g_A f \succ f_A g$. So, $\hat{f}_A \hat{g} \succeq \hat{g}_A \hat{f} \succ \hat{g} \succ \hat{f}$, again contradicting (2). □

(3) implies (4): This case is analogous to weak decomposability implies (1). □

(4) implies (5): This case is analogous to (1) implies (2). □

(5) implies weak decomposability: This case is analogous to (2) implies (3). □

We have established that weak decomposability is equivalent to (1) through (5), so weak decomposability implies (6) since (6) is an immediate consequence of (1) and (4). By P6*, if $g \succ f_A g$, then there is an event $E \subset A$ such that $b_E f_{A\backslash E} g \sim g$. So if (6) is satisfied but (3) is violated, i.e., $g \succ f_A g$ and $g \succ g_A f$ but $f \succeq g$, we can apply this property (substituting b into f first on A and then on $S\backslash A$) to obtain an act, $\hat{f}$, such that $g \sim \hat{f}_A g$ and $g \sim g_A \hat{f}$. But, given P3, by construction, $\widehat{f} \succ f \succ g$, contradicting (6). ■

Proof of Proposition 3. We first show that the utility function V is uniquely defined, and therefore that $\succeq$ satisfies P1. For any act f in $\mathcal{F}$, let $\Phi_f(w) := \sum_x \varphi\left(x, f^{-1}(x), w\right)$. By properties 2 and 3, for all events A in $\mathcal{E}\backslash\mathcal{N}$, $\varphi(x, A, w) \geq \varphi(x, A, x) = 0$ if and only if $x \geq w$. By our choice of $\mathcal{X}$, $\underline{x} \leq x \leq \bar{x}$ for all x in $f(S)$. Therefore, $\Phi_f(\bar{x}) \leq 0 \leq \Phi_f(\underline{x})$. So, by property 1 and 2, there exists a unique w $(=: V(f))$ with $\Phi_f(w) = 0$. For any constant act x, we have $V(x) = x$, hence property 2 implies that $\succeq$ satisfies P3.

We next show that, if $\mathcal{S}$, $\mathcal{E}$, and φ satisfy properties 2, 3 and 4, then φ is state monotonic; that is, for all pairs of events A and B in $\mathcal{E}$, and all x and w in $\mathcal{X}$, if $A \subset B$ then $|\varphi(x, A, w)| \leq |\varphi(x, B, w)|$. To see this, fix $A \subset B$ and fix x and w. By the definition of $\mathcal{N}$ and property 3, if $x = w$ then $\varphi(x, A, w) = \varphi(x, B, w) = 0$. By the definition of $\mathcal{N}$ and properties 2 and 3, if $x > w$ then $\varphi(x, E, w) \geq 0$ for all events E in $\mathcal{E}$ (with strict inequality for E in $\mathcal{E}\backslash\mathcal{N}$). So, in particular, $\varphi(x, A, w) \geq 0$ and $\varphi(x, B\backslash A, w) \geq 0$. Property 4 then gives $\varphi(x, A, w) \leq \varphi(x, B, w)$. The case for $x < w$ is similar.

Now, for P6, fix two acts f and g with $f \succ g$, and an outcome $\hat{x}$ in $\mathcal{X}$. Notice that $f \succ g$ implies that there exists an $\varepsilon > 0$, and a utility level w in $[\underline{x}, \bar{x}]$, for which $\Phi_f(w) > \varepsilon$ and $-\varepsilon > \Phi_g(w)$. Let T be the union of the range of f and g, which is a finite set. For each $x \in T \cup \{\hat{x}\}$, use property 5 to find a finite partition $\mathcal{P}_x$ such that $|\varphi(x, A, w)| < \frac{1}{1+\#(T\cup\{\hat{x}\})}\varepsilon$ for any $A \in \mathcal{P}_x$. Let $\mathcal{P}$ be the coarsest common refinement of $\{\mathcal{P}_x : x \in T \cup \{\hat{x}\}\}$. $\mathcal{P}$ is a finite partition. Pick any A in $\mathcal{P}$. By property 4, $\Phi_f(w) - \Phi_{\hat{x}_A f}(w) = \left(\sum_{\{x: A\cap f^{-1}(x)\neq\emptyset\}} \varphi\left(x, f^{-1}(x) \cap A, w\right)\right)$ $-\varphi(\hat{x}, A, w)$. By state monotonicity, $\varphi\left(x, f^{-1}(x) \cap A, w\right) < \frac{1}{1+\#(T\cup\{\hat{x}\})}\varepsilon$ for each x with $A \cap f^{-1}(x) \neq \emptyset$. Since the set $\left\{x : A \cap f^{-1}(x) \neq \emptyset\right\}$ has at most $\#(T \cup \{\hat{x}\})$ elements, we have $\left|\varphi_f(w) - \varphi_{\hat{x}_A f}(w)\right| < \varepsilon$. Similarity, $\left|\varphi_g(w) - \varphi_{\hat{x}_A g}(w)\right| < \varepsilon$ for each A in $\mathcal{P}$. These in turn imply $\varphi_{\hat{x}_A f}(w) > 0$ and $0 > \varphi_{\hat{x}_A g}(w)$. Thus $\hat{x}_A f \succ g$ and $f \succ \hat{x}_A g$ as required. Continuity with respect to non-null partitions follows immediately from property 1 (continuity with respect to the first and third arguments).

Finally, for weak decomposability, suppose $g_A f \succ f$ and $f_A g \succ f$. Then,
$\sum_x \varphi\left(x, g^{-1}(x) \cap A, V(f)\right) + \sum_x \varphi\left(x, f^{-1}(x) \backslash A, V(f)\right) > 0$ and
$\sum_x \varphi\left(x, f^{-1}(x) \cap A, V(f)\right) + \sum_x \varphi\left(x, g^{-1}(x) \backslash A, V(f)\right) > 0$. So
$\sum_x \varphi\left(x, g^{-1}(x) \cap A, V(f)\right) + \sum_x \varphi\left(x, g^{-1}(x) \backslash A, V(f)\right) > 0$, which implies $g \succ f$. ■

Proof of Proposition 4. Parts of the argument adapt ideas of Chew & Epstein (1989) and Wakker & Zank (1996). We use the following standard result.

Fact 1 Let $X = \Pi_{i=k}^{K} X_k$ with $K > 1$ and where X_k is an open interval in $\mathbb{R}$ for every k. Suppose U and V are continuous, additively separable on X, i.e., $U(\mathrm{x}) = \sum_{k=1}^{K} u_k(x_k)$ and $V(\mathrm{x}) = \sum_{k=1}^{K} v_k(x_k)$, and $U(\mathrm{x}) \geq U(\mathrm{y})$ iff $V(\mathrm{x}) \geq V(\mathrm{y})$. Then there is a unique set of constants $a, b_1, \ldots, b_K$ with $a > 0$, such that $u_k = av_k + b_k$ for all k.

Fix a non null partition $\mathcal{A} = \{A_k : k = 1, \ldots, K\}$ with $K > 3$, which exists by P6. Each act in $\mathcal{F}^{\mathcal{A}}$ can be regarded as an element of $\mathbb{R}^K$, so we naturally write an act f as $\mathrm{x} := (x_1, \ldots, x_K)$ where $x_k = f(A_k)$. We use the convention of writing f_{-k} or x_{-k} for the vector (act) that is obtained by dropping the kth element of f, and we write (x_{-k}, a) for the vector where kth coordinate is a and the other elements are given by the corresponding elements of vector x. Similarly, we write $(\mathrm{x}_{-(k,k')}, a, b)$ for the vector where kth and k'th coordinates are replaced with a and b, respectively.

For each outcome z in $(\underline{x}, \bar{x})$, and for each coordinate k, define $\mathcal{I}(z) = \{f \in \mathcal{F}^{\mathcal{A}} : f \sim z\}$, and $\mathcal{I}_{-k}(z) = \{\mathrm{x}_{-k} \in \mathbb{R}^{K-1} : \exists f \in \mathcal{I}(z), f_{-k} = \mathrm{x}_{-k}\}$. By construction, $(z, \ldots, z) \in \mathcal{I}_{-k}(z)$, so $\mathcal{I}_{-k}(z)$ is a non-empty set. By continuity, $\mathcal{I}_{-k}(z)$ is a closed subset of $[\underline{x}, \bar{x}]^{K-1}$ hence it is compact. Since $z \in (\underline{x}, \bar{x})$, again by continuity, the set $\mathcal{I}_{-k}(z)$ has a non-empty interior. It can be readily checked that $\mathcal{I}_{-k}(z)$ is the closure of its interior, where the boundary points lie in sets of the form $[\mathcal{I}_{-k}]_i(z) := \{\mathrm{x}_{-k} \in \mathcal{I}_{-k}(z) : \mathrm{y}_{-k} \in \mathcal{I}_{-k}(z) \Rightarrow (\mathrm{y}_{-k})_i \geq (\mathrm{x}_{-k})_i\}$, and $[\mathcal{I}_{-k}]^i(z) = \{\mathrm{x}_{-k} \in \mathcal{I}_{-k}(z) : \mathrm{y}_{-k} \in \mathcal{I}_{-k}(z) \Rightarrow (\mathrm{y}_{-k})_i \leq (\mathrm{x}_{-k})_i\}$, where $(\mathrm{y}_{-k})_i$ and $(\mathrm{x}_{-k})_i$ are the i th element of y_{-k} and x_{-k}, respectively.

Let $c_k(.; z) : \mathcal{I}_{-k}(z) \to \mathcal{X}$ be defined by the rule:

$$(\mathrm{x}_{-k}, c_k(\mathrm{x}_{-k}; z)) \sim z.$$

The function $c_k(\mathrm{x}_{-k}; z)$ is well defined by P1 (order) , P3 (eventwise monotonicity), continuity and the construction of $\mathcal{I}_{-k}(z)$. Define a binary relation $\succeq_k^z$ on $\mathcal{I}_{-k}(z)$ by the rule:

$$\mathrm{x}_{-k} \succeq_k^z \mathrm{y}_{-k} \;\Leftrightarrow\; c_k(\mathrm{x}_{-k}; z) \leq c_k(\mathrm{y}_{-k}; z).$$

So $\succeq_k^z$ is a well defined continuous relation on $\mathcal{I}_{-k}(z)$. Since each A_k in $\mathcal{A}$ is non null, this preference relation is strictly monotonic by P3. From Corollary 2 in Section 2, weak decomposability implies that each such relation $\succeq_k^z$ satisfies Debreu (1959)'s separability condition. We claim that each $\succeq_k^z$ admits a continuous, additively separable utility representation $U^k(\cdot; z)$ on $\mathcal{I}_{-k}(z)$ by Segal (1992, Theorem 1 and 2). We shall argue that all the conditions of his theorems are satisfied. The relation $\succeq_k^z$ is continuous and strictly monotone.

We have seen that the domain $\mathcal{I}_{-k}(z)$ is compact and equals the closure of its interior. The "richness of boundary" condition (the fifth condition in Theorem 2) is satisfied since, by continuity, each $[\mathcal{I}_{-k}]_i(z)$ and $[\mathcal{I}_{-k}]^i(z)$ is have non-empty relative interior.

Since the relation $\succeq$ is continuous and strictly monotonic on $\mathcal{X}^K$, each of its indifference surfaces is connected (in fact arc-connected) subset of $\mathbb{R}^K$. Thus the domain $\mathcal{I}_{-k}(z)$ is connected since it is the projected image of a connected set. To show that each truncation of $\mathcal{I}_{-k}(z)$ (Segal's $S(i,c)$) and each indifferent surface of $\succeq_k^z$ is connected, we can apply the argument of Lemma in Chew, Epstein & Wakker (1993, page 184).

For each k write $U^k(\mathbf{x}_{-k};z) = \sum_{i\neq k} u_i^k(x_i;z)$, where, by construction and P3, for each k, u_i^k is increasing and continuous in x_i. We can normalize each $u_i^k(z;z) = 0$ for every $i \neq k$. By construction, for each k and k' with $k \neq k'$, and each $i \notin \{k,k'\}$, the domains of the functions $u_i^k(\cdot;z)$ and $u_i^{k'}(\cdot;z)$ are the same. The result below shows that these functions are essentially the same.

Lemma A *For any k and k' with $k \neq k'$, and for any $i \notin \{k,k'\}$, we can find a unique positive constant $\beta_{k'}^k$ such that $\beta_{k'}^k u_i^k(\cdot;z) = u_i^{k'}(\cdot;z)$. Moreover, if k, l, and m are distinct, then $\beta_l^k \beta_m^l = \beta_m^k$.*

Proof. Without loss of generality, set $k := K-1$, $k' := K$. Consider the set $\mathcal{I}_{-(k,k')}(z) = \{\mathbf{x} \in \mathbb{R}^{K-2} : \exists a \in \mathcal{X}, \text{ such that } (\mathbf{x},a,a) \sim z\}$. By continutity this set has a non-empty interior. For each $i \notin \{K-1,K\}$, if x_i is in the domain of $u_i^K(\cdot;z)$ then there exists a vector $\mathbf{x}$ in $\mathbb{R}^{K-2}$ such that $(\mathbf{x}_{-i},x_i) \in \mathcal{I}_{-(k,k')}(z)$. Similarly, for each x_i is in the domain of $u_i^{K-1}(\cdot;z)$. We shall show that the functions $\sum_{i\neq k,k'} u_i^k(x_i;z)$ and $\sum_{i\neq k,k'} u_i^{k'}(x_i;z)$ induce the same preordering on $\mathcal{I}_{-(k,k')}(z)$. For any $\mathbf{x}_{-(K,K-1)}$, $\mathbf{y}_{-(K,K-1)}$ in $\mathcal{I}_{-(k,k')}(z)$, there are outcomes a and b such that $(\mathbf{x}_{-(K,K-1)},a,a) \sim (\mathbf{y}_{-(K,K-1)},b,b) \sim z$. Thus by construction, for any $\mathbf{x}_{-(K,K-1)}$, $\mathbf{y}_{-(K,K-1)}$, we have: $\sum_{i\neq k,k'} u_i^k(x_i;z) \geq \sum_{i\neq k,k'} u_i^k(y_i;z)$ $\Leftrightarrow U^k((\mathbf{x}_{-(K,K-1)},a,a)_{-k};z) \geq U^k((\mathbf{y}_{-(K,K-1)},a,a)_{-k};z) \Leftrightarrow a \leq b$ (by the construction of $\succeq_k^z$) $\Leftrightarrow U^{k'}((\mathbf{x}_{-(K,K-1)},a,a)_{-k'};z) \geq U^{k'}((\mathbf{y}_{-(K,K-1)},a,a)_{-k'};z) \Leftrightarrow \sum_{i\neq k,k'} u_i^{k'}(x_i;z) \geq \sum_{i\neq k,k'} u_i^{k'}(y_i;z)$. By Fact 1, with our normalization $u_i^k(z;z) = 0$, for any k and k' with $k \neq k'$, we can find a unique positive constant $\beta_{k'}^k$ such that $\beta_{k'}^k u_i^k(\cdot;z) = u_i^{k'}(\cdot;z)$ for any $i \notin \{k,k'\}$. To see the second half of the claim, since $K \geq 4$, pick an i that is distinct from k,l, or m. We have $u_i^m(\cdot;z) = \beta_m^l\left(u_i^l(\cdot;z)\right) = \beta_m^l\left(\beta_l^k u_i^k(\cdot;z)\right)$, and $u_i^m(\cdot;z) = \beta_m^k\left(u_i^k(\cdot;z)\right)$. So $\beta_l^k\beta_m^l = \beta_m^k$, as required. □

Now construct 'utility' functions ξ_k, $k = 1,\ldots,K$, by the rule: $\xi_1(\cdot;z) = \beta_1^K u_1^K(\cdot;z)$, and $\xi_i(\cdot;z) = u_i^1(\cdot;z)$ for $i > 1$. Write $\Xi(\mathbf{x};z) = \sum_k \xi_k(x_k;z)$.

Lemma B *If $\mathbf{x}$ and $\mathbf{y}$, both in $\mathbb{R}^K$, have a common component and $\mathbf{x} \sim \mathbf{y} \sim z$, then*

$\Xi(\mathbf{x}; z) = \Xi(\mathbf{y}; z)$.

Proof. Let $\mathbf{x} \sim \mathbf{y} \sim z$ and suppose we have $\Xi(\mathbf{x}; z) > \Xi(\mathbf{y}; z)$ while $\xi_k(x_k; z) = \xi_k(y_k; z)$ for some k. Consider two cases. If $\xi_1(x_1; z) = \xi_1(y_1; z)$, then $\Xi(\mathbf{x}; z) > \Xi(\mathbf{y}; z)$ implies $u_2^1(x_2; z) + \cdots + u_K^1(x_K; z) > u_2^1(y_2; z) + \cdots + u_K^1(y_K; z)$, but then $c_1(\mathbf{x}_{-1}; z) < c_1(\mathbf{y}_{-1}; z)$ must hold by the construction of U^1. But, since $\mathbf{x} \sim \mathbf{y} \sim z$, $c_1(\mathbf{x}_{-1}; z) = x_1$ and $c_1(\mathbf{y}_{-1}; z) = y_1$: a contradiction. In the second case $\xi_k(x_k; z) = \xi_k(y_k; z)$ for some $k > 1$. Then, by the definition of ξ_i, $\beta_k^1 \sum_{i \neq k} \xi_i(x_i; z) = \beta_k^1 \beta_1^K u_1^K(x_1; z) + \sum_{i \neq 1,k} \beta_k^1 u_i^1(x_i; z) = \sum_{i \neq k} u_i^k(x_i; z)$. So, $\Xi(\mathbf{x}; z) > \Xi(\mathbf{y}; z)$ implies $\sum_{i \neq k} u_i^k(x_i; z) > \sum_{i \neq k} u_i^k(y_i; z)$, which implies $c_k(\mathbf{x}_{-k}; z) < c_k(\mathbf{y}_{-k}; z)$. But, since $\mathbf{x} \sim \mathbf{y} \sim z$, $c_k(\mathbf{x}_{-k}; z) = x_k$ and $c_k(\mathbf{y}_{-k}; z) = y_k$: a contradiction. □

In the proof of the Lemma, for fixed z, we showed that for all k, $\beta_k^1 \sum_{i \neq k} \xi_i(x_i; z) = \sum_{i \neq k} u_i^k(x_i; z)$ (where $\beta_1^1 = 1$). We use this trick again below.

Lemma C *For any* $\mathbf{x}$ *in* $\mathbb{R}^K$, $\mathbf{x} \sim z$ *holds if and only if* $\Xi(\mathbf{x}; z) = 0$.

Proof. It suffices to show that $(x_1, \ldots, x_K) \sim z$ if and only if $\Xi(\mathbf{x}; z) = 0$. Suppose $\mathbf{x} = (x_1, \ldots, x_K) \sim z$, but $\Xi(\mathbf{x}; z) > 0$. We first claim that $\xi_k(x_k; z) > 0$ for every k is impossible. To see this, similar to before, $\sum_{k=2}^K \xi_k(x_k; z) > 0$ implies $c_1(\mathbf{x}_{-1}; z) < z$. But since $\mathbf{x} \sim z$, we have $c_1(\mathbf{x}_{-1}; z) = x_1$. So, by monotonicity, $\xi_1(x_1; z) < \xi_1(z; z) = 0$: a contradiction. The same argument rules out $\xi_k(x_k; z) < 0$ for every k. So, we can assume $\xi_k(x_k; z) > 0$ and $\xi_{k'}(x_{k'}; z) < 0$ for some k and k'.

Now $\xi_k(x_k; z) > 0$ implies $x_k > z$. If $\sum_{i \neq k} \xi_i(x_i; z) \geq 0$, then, by the same reasoning as the previous lemma, we get $\sum_{i \neq k} u_i^k(x_i; z) \geq \sum_{i \neq k} u_i^k(z_i; z)$, so $\mathbf{x}_{-k} \succeq_k^z z_{-k}$. But, since $\mathbf{x} \sim z$, this implies $x_k \leq z$: a contradiction. So $\sum_{i \neq k} \xi_i(x_i; z) < 0$ must hold. Hence, we have $\Xi(\mathbf{x}; z) = \Xi\left((\mathbf{x}_{-(k,k')}, x_k, x_{k'}); z\right) > 0 > \sum_{i \neq k} \xi_i(x_i; z) + \xi_k(z; z) = \Xi((\mathbf{x}_{-(k,k')}, z, x_{k'}); z)$.

We claim that there is an outcome a^1 with $\xi_k(x_k; z) > \xi_k(a^1; z) > \xi_k(z; z) = 0$ and such that $\Xi\left((\mathbf{x}_{-k}, a^1); z\right) = 0$. To see this, note by construction that for all $y_k \in (z, x_k)$, $(\mathbf{x}_{-(k,k')}, z, x_{k'}) \prec (\mathbf{x}_{-(k,k')}, y_k, x_{k'}) \prec z \preceq (\mathbf{x}_{-(k,k')}, x_k, z)$. Hence, by the continuity of $\succeq$, there exists $y_{k'} \in (x_{k'}, z)$ such that $(\mathbf{x}_{-(k,k')}, y_k, y_{k'}) \sim z$. Therefore the domain of $\xi_k(\cdot; z)$ contains $[z, x_k]$. Since $\xi_k(\cdot; z)$ is continuous, the claim follows by the intermediate value theorem. Set $\mathbf{w}^1 := (\mathbf{x}_{-k}, a^1)$. By construction, $x_k > a^1 > z$, and so $\mathbf{w}^1 \prec \mathbf{x} \sim z$.

We claim there exists $b^1 \in (x_{k'}, z]$, such that $(\mathbf{x}_{-(k,k')}, a^1, b^1) \sim z$. To establish this, it is enough to show $(\mathbf{x}_{-(k,k')}, a^1, z) \succeq z$. Suppose on the contrary, $z \succ (\mathbf{x}_{-(k,k')}, a^1, z)$. Then there exists $a \in (a^1, x_k]$ such that $(\mathbf{x}_{-(k,k')}, a, z) \sim z$, since $(\mathbf{x}_{-(k,k')}, x_k, z) \succeq z$. Furthermore, since $\Xi((\mathbf{x}_{-k}, a); z) = 0$ so $\Xi\left((\mathbf{x}_{-(k,k')}, a, z); z\right) > 0$, that is, $\sum_{i \neq k,k'} \xi_i(x_i; z) + \xi_k(a; z) > 0$. Hence, we have $c_{k'}((\mathbf{x}_{-(k,k')}, a); z) < z$, which contradicts $\left(\mathbf{x}_{-\{k,k'\}}, a, z\right) \sim z$.

To summarize: We started with $\mathbf{x} \sim z$, but $\Xi(\mathbf{x}; z) > 0$. We have $x_k > a^1 > z \geq b^1 > x_{k'}$; and we have constructed a vector $\mathbf{w}^1 := (\mathbf{x}_{-(k,k')}, a^1, x_{k'}) \prec z$ but with $\Xi(\mathbf{w}^1; z) = 0$, and a vector $\mathbf{x}^1 = (\mathbf{x}_{-(k,k')}, a^1, b^1) \sim z$, but (like the original $\mathbf{x}$) with $\Xi(\mathbf{x}^1; z) > 0$. So by the same construction, we can find an a^2 with $a^1 > a^2 > z$, and a b^2 with $z \geq b^2 > b^1$. So if $b^1 = z$, we would have obtained a contradiction. As before, let $\mathbf{w}^2 := (\mathbf{x}_{-(k,k')}, a^2, b^1) \prec z$ but with $\Xi(\mathbf{w}^2; z) = 0$, and $\mathbf{x}^2 := (\mathbf{x}_{-(k,k')}, a^2, b^2) \sim z$ but with $\Xi(\mathbf{x}^2; z) > 0$. Repeating this process, we obtain sequences $\{\mathbf{x}^n : n = 1, \ldots\}$ and $\{\mathbf{w}^n : n = 1, \ldots\}$, where $\Xi(\mathbf{w}^n; z) = 0$ and $\mathbf{x}^n \sim z$ for all n, and their kth and k'th components a^n and b^n constitute monotone bounded sequences. Thus both sequences converge, and $\lim \mathbf{x}^n$ and $\lim \mathbf{w}^n$ must be the same by construction. Let $\bar{\mathbf{x}}$ be the common limit point. By continuity, $\lim \mathbf{x}^n = \bar{\mathbf{x}} \sim z \sim \mathbf{x}$, and, by the continuity of Ξ, $\lim \Xi(\mathbf{w}^n; z) = \Xi(\bar{\mathbf{x}}; z) = 0$. Since the $K-2$ unchanged components of $\bar{\mathbf{x}}$ are equal to those of $\mathbf{x}$, we must have $\Xi(\bar{\mathbf{x}}; z) = \Xi(\mathbf{x}; z)$ by the previous lemma, but this contradicts $\Xi(\mathbf{x}; z) > 0$. An analogous argument shows that it is also impossible to have $\mathbf{x} \sim z$ and $\Xi(\mathbf{x}; z) < 0$. So $x \sim z$ implies $\Xi(\mathbf{x}; z) = 0$.

Conversely, suppose $\Xi(\mathbf{x}; z) = 0$ but $z \succ \mathbf{x}$. Then we can start with $\mathbf{w}^1$ in the construction of the sequences used above to obtain a contradiction. Similarly, $\Xi(x; z) = 0$ but $x \succ z$ is also impossible. □

We have shown that $\Xi(\mathbf{x}; z) = 0$ holds if and only if $\mathbf{x} \sim z$. For each $\mathbf{x}$ in $\mathbb{R}^K$, define $V(\mathbf{x}) = z$ such that z is the (unique) outcome indifferent to $\mathbf{x}$. The function V is continuous, and is given by the rule $\Xi(\mathbf{x}, V(\mathbf{x})) \equiv \sum_k \xi_k(x_k, V(\mathbf{x})) = 0$.

Now we are ready to construct φ. Fix a non-null partition $\mathcal{A}$ with four elements as a reference point, and construct the function $\Xi^{\mathcal{A}}$ that is associated with $\mathcal{A}$ as above. With slight abuse of notation, write $\Xi^{\mathcal{A}}(f; z)$ for $\Xi^{\mathcal{A}}(\mathbf{x}; z)$ where $\mathbf{x}$ is the vector associated with the act f measurable with respect to $\mathcal{A}$. For each non-null event E, let $\mathcal{A}(E)$ be the partition generated by $\mathcal{A}$ and $\{E\}$, and construct $\Xi^{\mathcal{A}(E)}$ as above for $\mathcal{A}(E)$. By construction, $\Xi^{\mathcal{A}(E)}$ induces the same preference relation as $\Xi^{\mathcal{A}}$ on acts measurable with respect to $\mathcal{A}$. So, Fact 1, we can normalize $\Xi^{\mathcal{A}(E)}$ by a unique positive scalar for each z in such a way that $\Xi^{\mathcal{A}(E)}(f; z) = \Xi^{\mathcal{A}}(f, z)$ for any f that is measurable with respect to $\mathcal{A}$. So, by setting $\varphi(x, E, z) = \Xi^{\mathcal{A}(E)}(x_E z; z)$ if E is non-null, and $\varphi(x, E, z) = 0$ otherwise, we have constructed a well-defined function on $\mathcal{X} \times \mathcal{E} \times \mathbb{R}$.

To see φ has the desired properties, take any simple act f and let $\mathcal{B} := \{E_1, \ldots, E_K\}$ be the coarsest partition of S containing $\mathcal{A}$, for which f is measurable. Construct $\Xi^{\mathcal{B}}$ and normalize it as above. Then, for any E_i in $\mathcal{B}$, by construction, $\Xi^{\mathcal{B}}(x_{E_i} z; z) = \Xi^{\mathcal{A}(E_i)}(x_{E_i} z; z)$, so $\varphi(x, E, z) = \Xi^{\mathcal{B}}(x_E z; z)$. Notice that $\Xi^{\mathcal{B}}(x_{E_i} z; z) = \xi_i^{\mathcal{B}}(x, z)$ if E_i is non-null. Therefore, $\sum_k \varphi(f(E_k), E_k, z) = 0$ if and only if $\sum_k \xi_k^{\mathcal{B}}(f(E_k); z) = 0$ which, by construction, is equivalent to $f \sim z$. Moreover $\varphi(x, E_i \cup E_j, z) = \xi_i^{\mathcal{B}}(x,z) + \xi_j^{\mathcal{B}}(x,z) = \varphi(x, E_i, z) + \varphi(x, E_j, z)$, so φ is state additive.

To see that φ satisfies event-continuity, fix x and w in $(\underline{x}, \bar{x})$ and take $x > w$ (the case for $x < w$ is similar). Suppose $\varphi(\underline{x}, \mathcal{S}, w) \leq \varphi(\bar{x}, \mathcal{S}, w)$ (an analogous argument holds when the inequality is reversed). Fix an $\varepsilon < \varphi(\underline{x}, \mathcal{S}, w)$. It follows from the continuity of φ in its first argument, the fact that φ in increasing in its first argument and the intermediate value theorem that there exists an outcome $y < w$ such that $\varphi(w, \mathcal{S}, w) = 0 = \varphi(y, A_k, w) + \varphi(y, \mathcal{S}\backslash A_k, w) + \varepsilon$. By P6 it follows that there exists a finite partition $\mathcal{P} = \{A_k : k = 1, \ldots, K\}$ such that $w \succ x_{A_k} y$ for all A_k in $\mathcal{P}$. That is, $0 > \varphi(x, A_k, w) + \varphi(y, \mathcal{S}\backslash A_k, w)$. Hence $\varphi(x, A_k, w) < \varepsilon$, for all A_k in $\mathcal{P}$, as required. ■

Proof of Proposition 5.

(1)⇒(2): By Proposition 1 part (4), weak decomposability is equivalent to: for any pair of acts f and g in $\mathcal{F}$, and any event A in $\mathcal{E}$, $g \succeq f_A g$ and $g \succeq g_A f$ imply $g \succeq f$. Since $F = H(\mathcal{H}, \mathcal{P})$, by construction, f^* is in F. So it is enough to show that for any act g in F, if $g \succeq h_A g$ for all acts h in $\mathcal{H}$ and all events A in $\mathcal{P}$, then $g \succeq f$ for all acts f in F. Fix an arbitrary f in F. Since $F = H(\mathcal{H}, \mathcal{P})$, for all events A in $\mathcal{P}$, there exists an act h in $\mathcal{H}$ such that $h_A g = f_A g$. Thus, $g \succeq f_A g$ for all events A in $\mathcal{P}$. The conclusion follows from an inductive argument similar to the proof of Observation 1. □

(2)⇒(1): assume that weak decomposability does not hold. Then there exists two acts f and g in $\mathcal{F}$, and an event A in $\mathcal{E}$, such that $g \succeq f_A g$ and $g \succeq g_A f$ but $f \succ g$. Consider the decision problem with $\mathcal{P} = \{A, \mathcal{S}\backslash A\}$ and $F = H(\{g, f\}, \mathcal{P})$. Our algorithm stops immediately yielding $f^* := g$. But g is not optimal since $f \succ g$. □

(1)⇒(3): Fix a finite partition $\mathcal{P} := \{A_1, \ldots, A_n\}$ and a pair of acts f, g in $\mathcal{F}$ with $f \sim g$ and $f \succeq h$, for all acts h in $H(\{f,g\}, \mathcal{P})$. We will proceed by induction. Suppose as an induction hypothesis first that f is indifferent to the act in which the first $k-1$ elements of the partition are determined by g and the remaining $n-k+1$ elements of the partition are determined by f. For $k = 1$, take the act to be f itself. So, the induction hypothesis holds for $k = 1$. Consider $k > 1$.

Suppose $g_{A_1 \cup \ldots \cup A_{k-1} \cup A_k} f \succ f_{A_k} g$. Then, by the contra-positive of Proposition 1 part (1) applied to the event A_k, either $g_{A_1 \cup \ldots \cup A_{k-1} \cup A_k} f \succ g_{A_1 \cup \ldots \cup A_{k-1}} f$ which is indifferent to f (by the induction hypothesis); or $g_{A_1 \cup \ldots \cup A_{k-1} \cup A_k} f \succ g$ which is indifferent to f (by hypothesis of DPS*). But, $g_{A_1 \cup \ldots \cup A_{k-1} \cup A_k} f \in H(\{f, g\}, \mathcal{P})$, so (by a hypothesis of DPS*), $f \succeq g_{A_1 \cup \ldots \cup A_{k-1} \cup A_k} f$: a contradiction. An analogous argument rules out $f_{A_k} g \succ g_{A_1 \cup \ldots \cup A_{k-1} \cup A_k} f$. Therefore we have $g_{A_1 \cup \ldots \cup A_{k-1} \cup A_k} f \sim f_{A_k} g$.

As $g \succeq g_{A_1 \cup \ldots \cup A_{k-1}} f$, it follows from the contra-positive of Proposition 1 part (3) applied to the event A_k that either $g_{A_1\cup...\cup A_{k-1}\cup A_k}f \succeq g_{A_1\cup...\cup A_{k-1}}f$ or $f_{A_k}g \succeq g_{A_1\cup...\cup A_{k-1}}f$. But, by the conclusion of the last paragraph, this implies $g_{A_1\cup...\cup A_{k-1}\cup A_k}f \succeq g_{A_1\cup...\cup A_{k-1}}f$. Combining we have $f \succeq$ (by hypothesis of DPS*) $g_{A_1\cup...\cup A_{k-1}\cup A_k}f \succeq g_{A_1\cup...\cup A_{k-1}}f \sim$ (induction hypothesis) f. So, $f \sim g_{A_1\cup...\cup A_{k-1}\cup A_k}f$. But since the ordering of the elements in the partition was arbitrary, we are done. □

(3)⇒(1): We will show: not weak decomposability implies not DPS*. Assume weak decomposability does not hold. That is, there exists a pair of acts f and g, and an event A such that $g_Af \succ f$, $f_Ag \succ f$ and yet $f \succeq g$. Let w be the worst outcome in the ranges of f and g. Applying P6* twice, there exists a (non-null) event $E^* \subseteq A$ such that $w_{E^*}g_{A\backslash E^*}f \sim f$ and there exists a (non-null) event $E^{**} \subseteq S\backslash A$ such that $w_{E^{**}}f_Ag \sim f$. Consider the four acts of the form $h = h^1_Ah^2$ where $h^i \in \{w_{E^*}g_{A\backslash E^*}f, w_{E^{**}}f_Ag\}$. By construction, we have $w_{E^*}g_{A\backslash E^*}f \sim w_{E^{**}}f_Ag$ and $w_{E^*}g_{A\backslash E^*}f \sim f$. The fourth such act $w_{E^*}w_{E^{**}}g \prec g$, by P3 (this relation must be strict given $g_Af \succ w_{E^*}g_{A\backslash E^*}f$). Thus, given $f \succeq g$, we have $w_{E^*}g_{A\backslash E^*}f \succ w_{E^*}w_{E^{**}}g$, violating DPS*. □

(1)⇔(4). Up to the interpretation of notation, the proof identical to (1)⇔(2). ■

Proof of Observation 3: If we let $\mu(A)$ denote the Lebesgue measure of any (measurable) event A, notice first that for any act f where neither $\mu\left(f^{-1}(1)\right)$ nor $\mu\left(f^{-1}(0)\right)$ is equal to one, we have $\int_0^1 \varphi(f(s), s, 0)\, ds = \int_0^{1/3} f(s)\, ds + \int_{1/3}^1 (1-a) f(s)\, ds > 0$, $\int_0^1 \varphi(f(s), s, 1)\, ds = \int_0^{1/3}(1-a) f(s)\, ds + \int_{1/3}^1 f(s)\, ds - 1 < 0$, and $\int_0^1 \varphi_v(f(s), s, v)\, ds < 0$. Hence for every act f there exists a unique v satisfying $\int_0^1 \varphi(f(s), s, v)\, ds = 0$.[20]

Further as this functional conforms with the form in Proposition 3 it follows that $\succeq$ satisfies weak decomposability. To see that $\succeq$ also satisfies P4 observe that for any pair of outcomes $x > y$ and measurable event A, it follows from the implicit definition of $V(f)$ that

$$\begin{aligned} &\mu(A\cap R)(x - V(x_Ay)) + \mu(A\cap(S\backslash R))(1-a)(x - V(x_Ay)) \\ &+\mu((S\backslash A)\cap R)(1-a)(y - V(x_Ay)) + \mu((S\backslash A)\cap(S\backslash R))(y - V(x_Ay)) \quad = \quad 0 \end{aligned}$$

Collecting terms, we have

$$\begin{aligned} & [1 - a(\mu(A\cap(S\backslash R)) + \mu((S\backslash A)\cap R))]\, V(x_Ay) \\ = \quad & [\mu(A\cap R) + \mu(A\cap(S\backslash R))(1-a)]\, x + [\mu((S\backslash A)\cap R)(1-a) + \mu((S\backslash A)\cap(S\backslash R))]\, y \end{aligned}$$

Noting that $V(y) = y$ and by subtracting

$$[1 - a(\mu(A\cap(S\backslash R)) + \mu((S\backslash A)\cap R))]\, V(y)$$

[20] For any act f for which $\mu\left(f^{-1}(1)\right)$ (respectively, $\mu\left(f^{-1}(0)\right)$) is equal to one, $V(f)$ is naturally set to 1 (respectively, 0).

from both sides yields

$$\begin{aligned}&[1-a(\mu(A\cap(\mathcal{S}\backslash R))+\mu((\mathcal{S}\backslash A)\cap R))][V(x_Ay)-V(y)]\\=\ &[\mu(A\cap R)+\mu(A\cap(\mathcal{S}\backslash R))(1-a)][x-y].\end{aligned}$$

Hence

$$V(x_Ay)-V(y)=\frac{[\mu(A\cap R)+\mu(A\cap(\mathcal{S}\backslash R))(1-a)]}{[1-a(\mu(A\cap(\mathcal{S}\backslash R))+\mu((\mathcal{S}\backslash A)\cap R))]}(x-y). \tag{4}$$

Given (4), for any pair of events A and B, $V(x_Ay)>V(x_By)$ (that is, the individual prefers 'betting on A' to 'betting on B') if and only if

$$\frac{[\mu(A\cap R)+\mu(A\cap(\mathcal{S}\backslash R))(1-a)]}{[1-a(\mu(A\cap(\mathcal{S}\backslash R))+\mu((\mathcal{S}\backslash A)\cap R))]}>\frac{[\mu(B\cap R)+\mu(B\cap(\mathcal{S}\backslash R))(1-a)]}{[1-a(\mu(B\cap(\mathcal{S}\backslash R))+\mu((\mathcal{S}\backslash B)\cap R))]}.$$

Since this inequality does not depend on the values of x or y, Savage's P4 follows. ■

Proof of Proposition 6 We note that under the maintained assumptions, the associated measure μ is strongly continuous; that is, for any $\varepsilon>0$, there exists a finite partition $\{E_1,...,E_K\}$ of $\mathcal{S}$ such that $\mu(E_k)<\varepsilon$ for every k. It is known that the range of a strongly continuous measure is convex, hence in particular, there is an event with probability $\frac{1}{2}$.

Lemma D *Let V be a continuous function on the set of lotteries $\mathcal{L}(\mathcal{X})$. If V satisfies the property that, for every pair of lotteries P and Q, $V(P)=V(\frac{1}{2}P+\frac{1}{2}Q)$ implies $V(P)\le$ (respectively, $\ge$) $V(Q)$ then V is quasi-convex (resp. quasi-concave).*

Proof. Suppose on the contrary that V is not quasi-convex. Then there exist lotteries $\bar{P}$ and $\bar{Q}$ and a weight α in $(0,1)$, such that $V(\bar{P})=V(\bar{Q})$ but $V(\bar{P})<V(\alpha\bar{P}+(1-\alpha)\bar{Q})$. Define a function v on $[0,1]$ by the rule $v(x)=V(x\bar{P}+(1-x)\bar{Q})$. By construction, v is continuous, so it attains its maximum value $\bar{v}$ on $[0,1]$. Let $w:=\sup\{x\in[0,1]:v(x)=\bar{v}\}$. By construction, $v(w)=\bar{v}>v(1)$. We say that $v(w)$ is a strict (respectivley, weak) local maximum if there is a $\bar{\delta}>0$ such that, for all x in $(w-\bar{\delta},w+\bar{\delta})$, $x\neq w$, $v(x)<v(w)$ (respectivley, $\le$). Similarly define strict and weak local minima.

Suppose $v(w)$ is a strict local maximum. Then, by continutity, there is a $\delta>0$, such that no x in $(w-4\delta,w+4\delta)$ is a weak local minimum. Therefore, if y is in $(w-4\delta,w)$ then for all x in (y,w), we have $v(y)<v(x)<v(w)$; and if y is in $(w,w+4\delta)$ then for all x in (w,y), we have $v(y)<v(x)<v(w)$. Without loss of generality, let $v(w-\delta)\le v(w+\delta)$. Let $\hat{x}\in[w-\delta,w)$ be such that $v(\hat{x})=v(w+\delta)$. Let $t=w+\delta-\hat{x}$. Set $\hat{y}:=(w+\delta)+t$. By construction $\hat{y}$ is in $(w,w+4\delta)$, so $v(\hat{y})<v(w+\delta)$. But, $w+\delta=\frac{1}{2}\hat{x}+\frac{1}{2}\hat{y}$: a contradiction.

Suppose, then, that $v(w)$ is not a strict local maximum. By construction, $w<1$, and $v(x)<v(w)$ for all x in $(w,1]$. Let $2\delta=1-w$. Since $v(w)$ is not a strict local maximum,

there exists an $\hat{x}$ in $(w-\delta, w)$ such that $v(\hat{x}) = v(w)$. Set $\hat{y} := w + (w - \hat{x})$. By construction, $\hat{y} < 1$, $v(\hat{y}) < v(w)$, and $w = \frac{1}{2}\hat{x} + \frac{1}{2}\hat{y}$: a contradiction. The argument for quasi-concavity is similar. □

We can now proceed to prove the proposition.

(i) $\Rightarrow$ *(ii)* Suppose that *(ii)* fails to hold. Then from the lemma there exists lotteries P and Q for which $V(P) = V\left(\frac{1}{2}P + \frac{1}{2}Q\right) > V(Q)$ or $V(P) > V\left(\frac{1}{2}P + \frac{1}{2}Q\right) = V(Q)$. Since the argument is symmetric, assume the former. As μ is non-atomic there exists an event A with $\mu(A) = 1/2$ and there exist two acts $\widehat{f}$ and $\widehat{g}$ which satisfy: $\mu \circ \widehat{f}^{-1}(x) = \mu\left(\widehat{f}^{-1}(x) \cap A\right)/\mu(A) = P(x)$ and $\mu \circ \widehat{g}^{-1}(x) = \mu\left(\widehat{g}^{-1}(x) \cap A\right)/\mu(A) = Q(x)$ for all x in $f(\mathcal{S}) \cup g(\mathcal{S})$. So by probabilistic sophistication $\widehat{f} \sim \widehat{f}_A\widehat{g} \sim \widehat{g}_A\widehat{f} \succ \widehat{g}$. But by setting $f := \widehat{f}$ and $g := \widehat{g}$ we have a violation of implication (6) of Proposition 1. □

(ii) $\Rightarrow$ *(i)* To show weak decomposability we require $g_A f \succ f$ and $f_A g \succ f$ to imply $g \succ f$. Notice first that neither A nor $\mathcal{S}\backslash A$ may be null. For E in $\{A, \mathcal{S}\backslash A\}$ let P_E, Q_E in $\mathcal{L}_0(\mathcal{X})$ denote the lotteries defined by the rule: $P_E(x) := \mu\left(f^{-1}(x) \cap E\right)/\mu(E)$ and $Q_E(x) := \mu\left(g^{-1}(x) \cap E\right)/\mu(E)$. Thus, $g_A f \succ f$ corresponds to $V(\mu(A)Q_A + (1-\mu(A))P_{\mathcal{S}\backslash A}) > V(\mu(A)P_A + (1-\mu(A))P_{\mathcal{S}\backslash A})$ and $f_A g \succ f$ to $V(\mu(A)P_A + (1-\mu(A))Q_{\mathcal{S}\backslash A}) > V(\mu(A)P_A + (1-\mu(A))P_{\mathcal{S}\backslash A})$. As V is quasi-concave in probability mixtures it follows that $V(\frac{1}{2}[\mu(A)P_A + (1-\mu(A))P_{\mathcal{S}\backslash A}] + \frac{1}{2}[\mu(A)Q_A + (1-\mu(A))Q_{\mathcal{S}\backslash A}]) > V(\mu(A)P_A + (1-\mu(A))P_{\mathcal{S}\backslash A})$. Applying the quasi-concavity again yields $V(\mu(A)Q_A + (1-\mu(A))Q_{\mathcal{S}\backslash A}) > V(\mu(A)P_A + (1-\mu(A))P_{\mathcal{S}\backslash A})$. This in turn implies by probabilistic sophistication that $g \succ f$, as required. ■

References

Bhaskara Rao, K.P.S. and M. Bhaskara Rao (1983): *Theory of Charges.* Academic Press, London.

Chew, S.H. (1983): "A Generalization of the Quasilinear Mean with Applications to the Measurement of Income Inequality and Decision Theory Resolving the Allais Paradox", *Econometrica*, 51, 1065-1092.

Chew, S.H. (1989): "Axiomatic Utility Theories with the Betweenness Property", *Annals of Operations Research*, 19, 273-298.

Chew, S.H. and L. Epstein (1989): "A Unifying Approach to Axiomatic Non-expected Utility Theories," *Journal of Economic Theory*, 49, 207-240.

Chew, S.H., L. Epstein and P. Wakker (1993) "A Unifying Approach to Axiomatic Non-expected Utility Theories: Correction and Comment", *Journal of Economic Theory*, 59, 183-88.

DEBREU, G. (1959): "Topological Methods in Cardinal Utility Theory", in *Mathematical Methods in the Social Sciences*, eds: Kenneth J. Arrow, Samuel Karlin, and Patrick Suppes, Stanford University Press.

DEKEL, E. (1986): "An Axiomatic Characterization of Preferences under Uncertainty: Weakening the Independence Axiom", *Journal of Economic Theory*, 40, 304-318.

ELLSBERG, D. (1961): "Risk, Ambiguity, and the Savage Axioms", *Quarterly Journal of Economics*, 75, 643-69.

EPSTEIN L. AND M. LE-BRETON (1993): "Dynamically Consistent Beliefs Must Be Bayesian", *Journal of Economic Theory*, 61, 1-22.

GILBOA, I. (1987): "Expected Utility with Purely Subjective Non-Additive Probabilities", *Journal of Mathematical Economics*, 16, 65-88.

GUL, F. (1991): "A Theory of Disappointment Aversion", *Econometrica*, 59, 667-686.

GRANT, S. (1995): "Subjective Probability without Monotonicity: or How Machina's Mom May also be Probabilistically Sophisticated", *Econometrica*, 63, 159-189.

GRANT, S., A. KAJII AND B. POLAK (1992): "Many Good Choice Axioms: When can many be treated as one", *Journal of Economic Theory*, 56, 313–337.

GUL, F. AND O. LANTTO (1990): "Betweenness Satisfying Preferences and Dynamic Choice", *Journal of Economic Theory*, 52, 162-177.

HAMMOND, P. (1988): "Consequentialist Foundations for Expected Utility," *Theory and Decision*, 25, 25-78.

HOBBES, THOMAS (1651): *Leviathan.* Fontana Library edition 1962, William Collins & Sons, London.

JOHNSEN, T.H. AND J.B. DONALDSON (1985): "The Structure of Intertemporal Preferences under Uncertainty and Time Consistent Plans," *Econometrica*, 53, 1451-1458.

MACHINA, M. (1983): "Generalized Expected Utility Analysis and the Nature of Observed Violations of the Independence Axiom," in *Foundations of Utility and Risk Theory with Applications*, eds: Bernt Stigum and Fred Wenstop. D. Reidel Publishing, Dordrecht, Holland, 263-95.

————(1989): "Dynamic Consistency and Non-Expected Utility Models of Choice under Uncertainty", in *Journal of Economic Literature*, 27, 1622-1668.

MACHINA, M. AND D. SCHMEIDLER (1992): "A More Robust Definition of Subjective Probability", *Econometrica*, 60, 745-780.

SAVAGE, L.J. (1972): *The Foundations of Statistics.* Dover Publications, New York.

Sarin, R. and P. Wakker (1997): "Consistency in Dynamic Choice Situations with Nonexpected Utility Models" mimeo.

Segal, U. (1992): "Additively Separable Representations on Non-Convex Sets", *Journal of Economic Theory*, 56, 89-99.

Schlee, E. (1997): "The Sure Thing Principle and the Value of Information", *Theory and Decision*, 42, 21-36.

Skiadas, C. (1997a): "Subjective Probability under Additive Aggregation of Conditional Preferences," *Journal of Economic Theory*, 76, 242-271.

——————(1997b): "Conditioning and Aggregation of Preferences," *Econometrica*, 65, 347-367.

Wakker, P. and H. Zank (1996): "State Dependent Expected Utility for Savage's State Space; Or: Bayesian Statistics without Prior Probabilities", working paper, CentER, University of Tilburg.