The 'Sure-Thing Principle' and the Optimality of Consistent Plans*

Simon Grant

Australian National University, Canberra, AUSTRALIA

simon.grant@anu.edu.au

Berend Roorda

University of Twente, Enschede, THE NETHERLANDS

b.roorda@utwente.nl

Jingni Yang

Australian National University, Canberra, AUSTRALIA

jingni.yang@anu.edu.au

October 25, 2022

Abstract

We consider a decision-maker's preferences over a class of decision trees involving the selection of an act from a menu conditional on receipt of a signal. Given the decision maker's contingent choices in each tree maximize her (potentially non-consequentialist) conditional preferences, we show every (dynamically) consistent plan will be (ex ante) optimal with respect to her (static) preferences over acts, if and only if those preferences exhibit a property introduced by Grant et al. (2000) as a way to operationalize Savage's (1954) extralogical *Sure-Thing Principle*.

JEL Classification: D80, D81

Key words: uncertainty, ambiguity, consistent planning, decomposability

*This paper was originally part of an earlier project entitled "Decomposable splits, coherent beliefs and the ex ante optimality of consistent plans". We thank audiences at Royal Holloway U. of London, ANU, Erasmus U. and SAET2021, as well as Jurgen Eichberger, Yoram Halevy, John Stachurski, Ron Stauber and especially, Peter Wakker for comments and suggestions.

1 Introduction

Denise Makim (hereafter referred to by her initials, DM) is considering whether to invest in a wind farm or a small modular (nuclear) reactor. As she may be unsure about the (future) returns from investing in a project, we model any uncertainty she faces with a state space and identify each investment opportunity by an *act*, formally a mapping from that state space to the space of (relevant) outcomes, in this particular instance, the space of conceivable financial returns.

She views the result of the upcoming Presidential election as relevant for the attractiveness of each project. Since she need not decide until after she has learned the result of the election, she can make her choice of project contingent on the result. Suppose she considers only two results are possible for the election: either the Republican Party's nominee wins or the Democratic Party's nominee wins.

Hence, there are four possible behaviors (courses of action) DM might take:

1. invest in the wind farm irrespective of which party's nominee wins the presidency;

2. invest in the small modular reactor in the event the Republican wins, otherwise invest in the wind farm;

3. invest in the small modular reactor irrespective of which party's nominee wins the presidency; and,

4. invest in the wind farm in the event the Republican wins, otherwise invest in the small modular reactor.

Notice as each behavior induces a unique mapping from the state space to the space of outcomes, that is, what we have referred to above as an act, it can be identified with that *act*.

We are interested in DM's preferences between and behavior within a class of problems that, like the one described above, involve a choice from some fixed (finite) menu of acts that she can condition on receipt of a signal. We refer to these problems as (decision) trees. A special subclass of trees are those where DM's course of action is predetermined because the menu from which she is to make her choice, comprises a single act.[1] We refer to such trees as *sole-option trees* and shall identify the restriction

[1] More precisely, the menu conditioned on each outcome of a signal, is a singleton. This includes the case in which DM only can choose, say, windfarm when the Democrat wins and reactor otherwise.

of her preferences to sole-option trees with her (static) preferences over acts. We refer to this restriction of her preferences as her *sole-option preferences*.[2]

If we suppose DM anticipates correctly what her behavior in each tree will be, then it is natural to require her preferences over trees be congruent with her preferences over the corresponding acts her behavior generates. This idea is usually referred to as *sophisticated choice* to distinguish it from *myopic choice* where she fails to anticipate correctly her behavior.

However, what DM actually does within a tree need not necessarily be the same thing as what she would choose were she able to commit to those choices ex ante. Informally, it is natural to say DM's behavior is ex ante optimal if her behavior within each tree is best with respect to her sole-option preferences over the set of acts that could have been achieved by any potential behavior in that tree. For example, if the order in the listing above of DM's four possible behaviors in her investment choice problem corresponded to a strict ranking by her sole-option preferences over the four acts they induced, then ex ante optimal behavior would require her to invest in the wind farm irrespective of which party's nominee wins the presidency.

To deem whether DM's behavior is (dynamically) consistent requires us to enrich our description of DM by defining her conditional preferences and then checking whether her behavior is guided by those conditional preferences. To avoid requiring her conditional preferences be consequentialist, we define them for each *information* event (that is, any event on which she can condition her choice of act in some tree) *as well as* for each conceivable *baseline* act. We interpret this relation as saying when DM would prefer one act to another act knowing the conditioning event had obtained and that the baseline act would have determined the outcome for all the states that lay outside this conditioning event.

Equipped with this collection of conditional preference relations, we say DM's behavior within a specific tree is (dynamically) consistent if for each signal realization, her behavior is guided by the conditional preference relational that corresponds to her knowing the event in which that signal realization obtains has occurred and with the baseline act set equal to the act her behavior in the tree induces. And, if this is the case for all trees, then we refer to DM as a *consistent planner.*

One (essentially, tautological) way to ensure DM's ex ante optimal behavior is (dynamically) consistent is to link the specification of her conditional preferences to

[2] Implicitly this rules out any framing effects of the kind discussed in Kahneman and Tversky (1981). Furthermore, it also precludes any *intrinsic* preference for information as introduced and analyzed in the context of choice under risk in Grant et al. (1998).

her sole-option preferences so that by construction she wants to follow through with a plan of action that was judged (ex ante) optimal by her sole-option preferences. In the sequel we refer to this as *resolute conditioning*. This notion is closely related to the dynamic consistency properties considered in Johnsen and Donaldson (1985), Machina (1989), McClennen (1990), Karni and Schmeidler (1991) and Halevy (2004).

In this paper, however, we consider under what circumstances does the (partial) converse hold. That is, we address the following question?

> *Given DM engages in resolute conditioning, when can we conclude if she is a consistent planner that her behavior must be ex ante optimal?*

The answer turns out to be when (and *only when*) her sole-option preferences satisfy *(weak) decomposability*, a property introduced by Grant et al. (2000) as a way to operationalize Savage's (1954) extralogical *Sure-Thing Principle* (STP). As Grant et al. note, decomposability is implied by, but does not imply, Savage's postulate **P2**, that he proposed as a way to operationalize the STP. Indeed, we shall see that if DM's sole-option preferences strictly rank the four acts, corresponding to the four possible behaviors in her investment choice problem, in the same order in which we listed them above, then even though that entails a violation of **P2**, that ranking is not inconsistent with her sole-option preferences satisfying decomposability.

In the next section we formally introduce our framework of decision trees as well as the characterization of a decision-maker in terms of her preferences across trees and her behavior within trees. This characterization of the DM is enriched in section 3 by the specification of her conditional preferences and the notion of a consistent plan in a tree corresponding to behavior within that tree that is guided by these conditional preferences. This is followed by the formal definition of resolute conditioning that links the DM's conditional preferences to her sole-option preferences in such a way as to ensure any ex optimal behavior is consistent. Our main result (Theorem 1) establishes that sole-option preferences satisfying decomposability is both necessary and sufficient to guarantee the converse, namely, the behavior of a consistent planner must be ex ante optimal. Focusing our attention on DMs whose sole-option preferences satisfy decomposability, Section 4 demonstrates the computational advantages this property affords us in generating behavior that is ex ante optimal.

For the last part of the paper we consider sole-option preferences that allow for a separation between a decision-maker's *beliefs* (over the likelihood of events) and *evaluations* (of random outcomes). We begin in section 5 by specifying a family of sole-option preferences that admit a representation that can be seen as an extension of Machina and Schmeidler's (1992) class of probabilistically sophisticated non-expected utility maximizers to a setting that allows for ambiguity in the sense that the uncertainty the DM associates with some events cannot be quantified by a precise probability. We refer to members of this family as *Generalized Expected Uncertain Utility* Maximizers as the representation, like the Expected Uncertain Utility model of Gul and Pesendorfer (2014) involves a probability defined on a rich set of events the DM deems unambiguous. Adopting the nomenclature of Gul and Pesendorfer, we refer to this probability as the DM's *prior*. We show this prior maps any act to a belief function (a generalization of a probability measure) defined over outcome-sets. So, analogous to Machina and Schmeidler (1992), the second component of the representation is a (mixture) continuous and (outcome-set) monotonic function defined on belief functions. The final section provides examples of GEUU maximizers that have appeared in the literature and our main result is that a GEUU maximizer satisfies decomposability if and only if the function over belief functions in its representation satisfies the betweenness property of Chew (1983) and Dekel (1986).

2 Decision Trees, Preferences and Behavior

We adopt Machina and Schmeidler's (1992) rendition of Savage's (1954) setting of uncertainty. Let Ω be the state space. For any pair of events $D\,, E \subseteq \Omega$, $D \backslash E$ denotes the set of elements that are in D but not in E. Hence $\Omega \backslash E$ denotes the complement of E. We refer to some events as *information events* and denote by $\mathcal{I}$ the set of information events (with generic elements R, $\widehat{R}$, R', et cetera). We assume $\mathcal{I}$ is a *mosaic*, a collection of events in which *(i)* $\Omega \in \mathcal{I}$; *(ii)* if $R \in \mathcal{I}$ then $\Omega \backslash R \in \mathcal{I}$; and, *(iii)* if $R_1\,, \ldots\,, R_n$ are all in $\mathcal{I}$ and form a partition of Ω, then for all $i, j \in \{1, \ldots, n\}$, $R_i \cup R_j \in \mathcal{I}$.[3]

An uncertain prospect or *act* is a mapping a in which for each state ω in Ω, $a(\omega)$ is the outcome from some set X that is associated with the act when the state ω is realized. Every act is *simple* in the sense that its image is a finite subset of X. Let A denote the set of acts. For any pair of acts a and a' in A and any information event $R \in \mathcal{I}$, we write $a_R a'$ for the act that agrees with a on R and with a' on $\Omega \backslash R$.

Our focus is on decision problems or *trees* like the one described in the introduction above, in which the DM selects an act from a (fixed) menu after learning the

[3] See Kopylov (2007) for further discussion and elaboration of the role mosaics play in characterizing the set of events that lend themselves to precise quantification by a probability.

realization of a signal. More formally, a signal is described by an $\mathcal{I}$-adapted (onto) function $\sigma\colon \Omega \to S$ ($|S| < \infty$) with the interpretation that the realization of the signal is s when the (information) event $\sigma^{-1}(s)$ (in $\mathcal{I}$) obtains. A menu $M \subset A$ (with $|M| < \infty$) corresponds to the set of acts from which the DM chooses after receipt of the signal. Denote by T (with generic element $\tau = \langle (S\,,\sigma)\;,M\rangle$) the collection of all such trees. Let $\succsim_T$ be a complete and transitive preference relation over such trees. We assume this preference relation guides a choice, made ex ante by the DM, among trees.

For example, let $\widehat{\tau} = \left\langle \left(\widehat{S}\,,\widehat{\sigma}\right)\;,\widehat{M}\right\rangle$ denote the decision tree described in the introduction. Recall the DM (in this case, Denise Makim) faced a choice between investing in a wind farm or a small modular reactor that she could make contingent on the result of the next presidential election. Let wf (respectively, smr) denote the act corresponding to her investing in the wind farm (respectively, small modular reactor), thus $\widehat{M} = \{wf, smr\}$. Let e (respectively, d) correspond to Denise being informed that the Republican (respectively, Democrat) candidate has won the presidency. Hence $\widehat{\sigma}$ is the mapping from states to $\widehat{S} = \{e\,, d\}$ in which $\widehat{\sigma}^{-1}(e)$ (respectively, $\widehat{\sigma}^{-1}(d)$) is the event the Republican (respectively, Democrat) candidate wins the election.[4]

Within a given tree the DM chooses an act at each decision node following each possible signal realization. Let b be a *behavior function* that, for each tree $\tau = \langle (S\,,\sigma)\;,M\rangle$ in T assigns an act from M to each signal realization s in S and let b^τ be the corresponding behavior in the specific tree τ with $b^\tau_s \in M$ denoting the corresponding act selected in that tree after she sees the specific signal realization s. Denote by B the set of all possible behaviors. With a slight abuse of our notation, we shall identify the behavior b^τ in a specific tree τ with an act in A, by setting $b^\tau(\omega) := b^\tau_s(\omega)$ whenever ω is in $\sigma^{-1}(s)$. Finally, let $B(\tau)$ ($\subset A$) denote the set of acts that can be generated by some behavior in the tree τ. More formally, we set

$$B(\tau) := \big\{\widehat{a} \in A\colon \text{for each } s \in S\,,\ \text{and every pair of states } \omega\,,\omega' \in \sigma^{-1}(s)\,,$$
$$\widehat{a}(\omega) = \widehat{a}(\omega') = a(\omega) \text{ for some } a \in M.\}$$

Returning to the tree $\widehat{\tau}$ described in the introduciton, $b^{\widehat{\tau}} = (b^{\widehat{\tau}}_e\,, b^{\widehat{\tau}}_d) \in \{wf\,, smr\} \times$

[4] Recall DM views the two events $\widehat{\sigma}^{-1}(e)$ and $\widehat{\sigma}^{-1}(d)$ as not only mutually exclusive but also exhaustive.

$\{wf\,, smr\}$ and so can be identified with the act it induces from the set

$$B(\widehat{\tau}) = \{wf\,, wf_{\sigma^{-1}(e)}smr\,, smr\,, smr_{\sigma^{-1}(e)}wf\}\,.$$

A special class of trees are *sole-option trees* in which the menu contains a single act. Since there is only one possible behavior within each sole-option tree, we identify the restriction of $\succsim_T$ to sole-option trees (which we denote by $\succsim$) with the DM's (static) preferences over acts.

In the introduction we argued, a DM who knows her behavior should only be considered *sophisticated*, if her preferences over trees reflects her sole-option preferences over the corresponding acts her behavior generates in each tree. The following formalizes this idea.

Definition 1 *A DM is* sophisticated *if for all pairs of trees τ and τ' in T:*

$$\tau \succsim_T \tau' \text{ if and only if } b^{\tau} \succsim b^{\tau'}\,.$$

On the other hand, for her behavior to be considered *ex ante optimal*, it should be the case that the act her behavior generates within each tree is best with respect to her sole-option preferences over the set of acts that could have been achieved by any potential behavior in that tree. The following formalizes this idea.

Definition 2 *A DM's behavior b is ex ante optimal if for each tree τ in T, $b^{\tau} \succsim a$, for all acts $a \in B(\tau)$.*

Recall in our investment choice example from the introduction, we observed that if

$$wf \succ smr_{\sigma^{-1}(e)}wf \succ smr \succ wf_{\sigma^{-1}(e)}smr\,,$$

then ex ante optimal behavior entails $b_e^{\widehat{\tau}} = b_d^{\widehat{\tau}} = wf$ (or equivalently, $b^{\widehat{\tau}} = wf$).

Combining these ideas, we may view a DM whose behavior is both sophisticated and ex ante optimal, as somebody who, when given a choice between two decision trees, expresses a preference for one over the other if and only if there exists an achievable act in the former that dominates (according to her sole-option preferences) all the acts that are achievable in the latter. That is, $\tau \succsim_T \tau'$ if and only if there exists an (achievable) act a in $B(\tau)$ for which $a \succsim a'$, for all acts a' in $B(\tau')$.

3 Consistent Plans, Resolute Choice and The Sure-Thing Principle

To analyze the (dynamic) consistency of the DM's behavior, we enrich our description of the DM by defining, for each information event $R \in \mathcal{I}$ and each act a in A, her conditional preference relation $\succsim_R^a$. That is, we shall allow for the conditioning to depend not only on the conditioning event R but (potentially) also on a "base-line" act a. The DM's behavior within a specific tree τ will be deemed a (dynamically) consistent plan if the act selected for each signal realization is optimal with respect to her conditional preferences with $\sigma^{-1}(s)$ as the conditioning event and the act generated by b^τ as the baseline act. If this is the case for all trees then we shall refer to such a DM as a *consistent planner*.[5]

Definition 3 (Consistent Planning) *The DM is a* consistent planner *if for all trees $\tau = \langle (S\,,\sigma)\ ,M\rangle$ in T, b^τ is a consistent plan, that is,*

$$b_s^\tau \succsim_{\sigma^{-1}(s)}^{b^\tau} a \textit{ for all } a \in M \textit{ and all } s \in S\,.$$

So, for example, if the DM is a consistent planner and her behavior $b^{\widehat{\tau}}$ in the investment choice problem $\widehat{\tau}$ described above, generates the act wf, then this entails

$$wf \succsim_{\sigma^{-1}(e)}^{wf} smr \text{ and } wf \succsim_{\sigma^{-1}(d)}^{wf} smr\,.$$

Machina (1989), McClennen (1990), and Halevy (2004) (along with others) have argued a natural way to define conditional preferences is to link them to the unconditional preferences as follows.

Axiom 1 (Resolute Conditioning) *The DM's conditional preferences are* resolute with respect to a baseline act *if for each information event $R \in \mathcal{I}$ and each act a in A: $a' \succsim_R^a a'' \iff a'_R a \succsim a''_R a$, for every pair of acts $a'\,, a'' \in A$.*

If the DM engages in resolute conditioning then we see that by construction she will want to follow through with any plan that was judged ex ante optimal by her sole-option preferences.

Lemma 1 *Suppose Axiom 1 holds. If the DM's behavior is ex ante optimal then she is a consistent planner.*

[5] Such behavior is consistent in the sense of Siniscalchi (2011) and Strotz (1955).

Proof. Fix a DM characterized by $\left\langle \succsim_T , b , (\succsim^a_R)^{a\in A}_{R\in\mathcal{I}} \right\rangle$. Suppose b is ex ante optimal. Then for each tree $\tau = \langle (S , \sigma) , M \rangle \in T$, $b^\tau \succsim a$ for every $a \in B(\tau)$. So in particular, for each $s \in S$, $b^\tau \succsim a_{\sigma^{-1}(s)} b^\tau$ for every $a \in M$. But as Axiom 1 holds, this means that $b^\tau_s \succsim^{b^\tau}_{\sigma^{-1}(s)} a$ for every $a \in M$. That is, b^τ is a consistent plan. ■

When does the (partial) converse hold? That is, given Axiom 1 holds, when is the behavior of a consistent planner, necessarily ex ante optimal? The answer appears to be when the DM satisfies Savage's Sure-Thing Principle which can be (informally) expressed as follows:

> *If the DM weakly prefers a to a', either knowing that the event R obtains, or knowing that the complement of the event R obtains, then she weakly prefers a to a' even when **not knowing** whether R has obtained or not.*

Since objects like "a if the event R obtains" are at best subacts, we require a means to extend these subacts to the entire state space. Savage does so in a way that leads to a *consequentialist* conditional preference relation $\succsim_R$ derived from $\succsim$ by setting $a \succsim_R a'$ if and only if $a_R a'' \succsim a'_R a''$ *for every act* a''.[6] Hence in order for these conditional preferences to be well-defined and well-behaved, he requires $\succsim$ satisfy his postulate **P2**.

Axiom 2 (P2) *For any pair of acts a and a' in A and any information event $R \in \mathcal{I}$:*

$$a \succsim a'_R a \implies a_R a' \succsim a' .$$

To see how this "operationalizes" his sure-thing principle, first notice that the two conditional preference statements $a \succsim_R a'$ and $a \succsim_{\Omega\setminus R} a'$ require by definition (taking $a'' = a$) that both $a \succsim a'_R a$ and $a \succsim a_R a'$, hold. Moreover, if $\succsim$ satisfies Axiom 2 (**P2**) then $a \succsim a'_R a$ implies $a_R a' \succsim a'$. Hence from the transitivity of $\succsim$ we have $a \succsim a'$, as required.

However, as Lemma 1 demonstrates, consistent planning is perfectly compatible with conditional preferences that are not consequentialist. And indeed, recalling the DM's ranking of the acts in $B(\widehat{\tau})$, namely,

$$wf \succ smr_{\sigma^{-1}(e)} wf \succ smr \succ wf_{\sigma^{-1}(e)} smr ,$$

[6] The relation $\succsim_R$ is consequentialist in the sense that the preference between any pair of acts a and a' conditional on R obtaining only depends on those outcomes that can arise for a and a' in states that lie in the event R (or equivalently, is not affected by what outcomes might have arisen for either a or a' in states that are not in R).

we see this ranking *violates* **P2**, since $smr \succsim wf_{\sigma^{-1}(e)}smr$ and yet we do not have $smr_{\sigma^{-1}(e)}wf \succsim wf$, as **P2** requires.

So instead, we propose interpreting the statement that a would be weakly preferred to a' if R were known to obtain, along the lines of Axiom 1 (Resolute Conditioning). That is, we set $a \succsim_R^a a'$ if and only if $a \succsim a'_R a$.[7] With our alternative interpretation, the consistency condition we propose for $\succsim$ is the *(weak) decomposability property* of Grant et al. (2000).[8]

Axiom 3 (Decomposability) *For any pair of acts a and a' in A and any information event $R \in \mathcal{I}$:*

$$a \succsim a'_R a \text{ and } a \succsim a_R a' \implies a \succsim a' .$$

To see how our alternative interpretation can also "operationalize" the sure-thing principle, notice that the two conditional preference statements $a \succsim_R^a a'$ and $a \succsim_{\Omega \setminus R}^a a'$ entail (by definition) the two unconditional preference statements $a \succsim a'_R a$ and $a \succsim a_R a'$. By applying Axiom 3 (Decomposability) we obtain $a \succsim a'$, as required.

If a DM's conditional preferences are resolute with respect to a baseline act and she is a consistent planner then it turns out that the property of decomposability is both necessary and sufficient for any consistent plan to be ex ante optimal. The result is comparable to Halevy (2004, Proposition 1), which characterizes the absence of speculative trade in terms of decomposability.

Theorem 1 *Consider a DM characterized by $\left\langle \succsim_T , b , (\succsim_R^a)_{R \in \mathcal{I}}^{a \in A} \right\rangle$. Suppose Axiom 1 (Resolute Conditioning) holds. Then the following are equivalent.*

1. *If the DM is a consistent planner then her behavior is ex ante optimal.*

2. *The DM's sole-option preferences satisfy Axiom 3 (Decomposability).*

Proof

1. implies 2.

We establish this holds by proving the contrapositive, namely, if Axiom 3 does not hold, then we can find a consistent plan that is not ex ante optimal. More specifically, for Axiom 3 not to hold, means there exists an information event R in $\mathcal{I}$ and two

[7] Notice that instead of requiring $a_R a'' \succsim a'_R a''$ hold for *every* act a'', we need it hold *only* for $a'' = a$.

[8] Grant et al. (2000) state decomposability in terms of strict preferences but describe equivalence with several non-strict criteria in Lemma 1.

acts a and a' in A, for which $a \succsim a'_R a$ and $a \succsim a_R a'$, however $a' \succ a$. So consider the tree $\widehat{\tau}$ with $\widehat{S} = \{1, 2\}$, $\widehat{\sigma}(\omega) = 1$ if $\omega \in R$ and $\widehat{\sigma}(\omega) = 2$ if $\omega \notin R$, and $\widehat{M} = \{a, a'\}$. A behavior b in which b^τ is ex ante optimal for all $\tau \neq \widehat{\tau}$ and for which $b_1^{\widehat{\tau}} = b_2^{\widehat{\tau}} = a$ is a consistent plan. However, the behavior $\widetilde{b}$ for which $\widetilde{b}^\tau = b^\tau$ for all $\tau \neq \widehat{\tau}$ and $\widetilde{b}_1^{\widehat{\tau}} = \widetilde{b}_2^{\widehat{\tau}} = a'$ is ex ante superior.

2. implies 1.

To establish the ex ante optimality of any behavior b in which each b^τ is a consistent plan, we first establish the following property for any partition adapted to $\mathcal{I}$:

> *If there is no strict incentive to move from one act toward another act on any single element of that partition, then there is no strict incentive to move from the former to the latter globally.*

The following lemma expresses this more formally.

Lemma 2 *Fix a finite partition of events of the state space $\{R_1, \ldots, R_n\}$, with $R_i \in \mathcal{I}$ for all $i =$ ‘$1, \ldots, n$. If $a \succsim \widehat{a}_{R_i} a$ for all $i = 1, \ldots, n$ then $a \succsim \widehat{a}$.*

Proof. We proceed by induction. Set $\widehat{a}^k := \widehat{a}_{R_1 \cup \ldots \cup R_k} a$, so that $\widehat{a}^1 = \widehat{a}_{R_1} a$ and $\widehat{a}^n = \widehat{a}$. As an induction hypothesis, suppose that $a \succsim \widehat{a}^k$. By assumption this hypothesis holds for $k = 1$. For any $k \in \{1, \ldots, n-1\}$ we have $a \succsim \widehat{a}^{k+1}_{R_{k+1}} a = \widehat{a}_{R_{k+1}} a$ (by hypothesis) and $a \succsim a_{R_{k+1}} \widehat{a}^{k+1} = \widehat{a}^k$ (by the induction hypothesis). Hence by applying Axiom 3 we have $a \succsim \widehat{a}^{k+1}$, as required. ∎

Since by definition, for each tree $\tau = \langle (S, \sigma), M \rangle$ in T, there is no strict incentive to move away from the act prescribed by b^τ for each event $\sigma^{-1}(s)$ in the partition $\{\sigma^{-1}(s')\}_{s' \in S}$, it follows from lemma 2 there can be no strict incentive to move away globally to any act that can be generated by any other possible behavior, that is, any act in $B(\tau)$. Hence b is ex ante optimal.

∎

4 Computing ex ante optimal behavior

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. Axiom 3 (decomposability) may also be thought of as a condition under which it is enough to check "local" necessity conditions to find an ex ante optimal plan for each tree. To see how, fix a tree $\tau = \langle (S\,, \sigma)\ , M \rangle$ in which $M = \{m^1, \cdots, m^k\}$ and $S = \{s^1, \cdots, s^n\}$. Consider the following algorithm, which is a version of coordinate ascent:[9]

(1) Set $a^0 := m^1$ and $a^1 := m^1$

(2) Set $j := 0$

(3) Set $j := j + 1$ and $i := 0$

(4) Set $i := i + 1$

(5) If $m^i_{\sigma^{-1}(s^j)} a^1 \succ a^1$ then set $a^1 := m^i_{\sigma^{-1}(s^j)} a^1$

(6) If $i < k$ then go to (4).

(7) If $j < n$ then go to (3).

(8) Stop if $a^0 = a^1$, otherwise set $a^0 := a^1$ and go to (2)

Since the problem is finite, we know that this algorithm will stop. We want to ensure that where-ever it stops is an optimum. A necessary condition for a particular plan a to be optimal is that there is no other plan in $B(\tau)$ that differs from a on only an event $\sigma^{-1}(s)$ for some s in S and which is better. This is what the algorithm checks: $b^\tau \succsim m_{\sigma^{-1}(s)} b^\tau$ for all m in M and all s in S. In a sense this is a local condition: it considers deviations from the candidate optimal act at only one 'place' at a time. Decomposability is equivalent to requiring that this local necessary condition is always globally sufficient. The byproduct of our result is that we provide a necessary and sufficient condition for the convergence of coordinate ascent algorithm to a global optimum.

The computational convenience of Axiom 3 also shows in testing whether a given candidate plan is optimal. A crude sufficiency check would involve checking the candidate against all $|B(\tau)| - 1\,(= |M|^{|S|} - 1\,)$ other feasible plans. Given decomposability, however, we need only check the local necessary condition for optimality. This involves only $(|M| - 1) \times |S|$ comparisons.

[9] Coordinate ascent (decent) algorithm to our attention. is a widely used algorithm in operations research and computer science, see for example Wright (2015) and Nesterov (2012) for discussions. We thank John Stachurski for bringing this to our attention.

5 Generalized Expected Uncertain Utility Maximization

Up to this point, the assumptions we have made about the DM's preferences over trees, and in particular her sole-option preferences, have not implied there exists any way to separate the DM's beliefs about the likelihood of (information) events obtaining and her evaluations of (random) outcomes. In this section, we introduce a class of sole-option preferences that affords such a separation. They can be seen to be an extension of Machina and Schmeidler's (1992) class of probabilistically sophisticated non-expected utility maximizers to a setting that allows for ambiguity in the sense that the uncertainty the DM associates with some events cannot be quantified by a precise probability. In the following section we investigate what additionally is implied if the sole-option preferences satisfy Axiom 3 (decomposability).

In the tradition of the voluminous literature initiated by Ellsberg (1961), our DM perceives there to be ambiguity arising as a result of her possessing only *incomplete information* about the underlying stochastic process that determines the resolution of the uncertainty she faces. In particular, this means she is not comfortable quantifying with a precise probability the uncertainty she associates with each and every event. However, we assume she is able to do so for the class of information events. In common with Gul and Pesendorfer's (2014) family of expected uncertain utility (EUU) maximizers, we assume her information is "rich" and "nice" enough to allow us to represent her beliefs about the likelihood of information events obtaining with a probability that we refer to as the DM's prior.

Definition 4 (Prior) *A prior μ is a countably-additive and convex-ranged probability defined on $\mathcal{I}$ which we now take to be a σ-algebra of (risky) events.*[10] *For each information event $\widehat{R} \in \mathcal{I}$, with $\mu(\widehat{R}) > 0$, let $\mu_{\widehat{R}}$ denote the updated probability conditional on $\widehat{R}$ obtaining. That is,*

$$\mu_R(R) = \frac{\mu(R \cap \widehat{R})}{\mu(\widehat{R})} \text{ for all } R \in \mathcal{I}\,.$$

Even though the uncertainty associated with events that lie outside of $\mathcal{I}$ cannot be

[10] Countable-additivity requires the probability of the union of a countable collection of disjoint elements from $\mathcal{I}$ equals the infinite sum of the probabilities of these events. For μ to be convex-ranged requires for any event R in $\mathcal{I}$ and any r in $(0,1)$ there exists a subset $R' \subset R$ that is in $\mathcal{I}$ and for which $\mu(R') = r\mu(R)$.

precisely quantified by μ, nevertheless μ can provide a lower estimate of the likelihood of such events obtaining and thus can be used to map each act a in A, to a (potentially non-additive) measure over sets of outcomes, which we denote by β_μ^a and define by setting for each $Y \subset X$,

$$\beta_\mu^a(Y) := \sup_{R \in \mathcal{I}\,, R \subseteq a^{-1}(Y)} \mu(R)$$

Intuitively, $\beta_\mu^a(Y)$ may be interpreted as a lower bound the prior μ places on the likelihood the act a will lead to an outcome in the set Y obtaining. For any act a that is measurable with respect to μ, that is, $a^{-1}(Y) \in \mathcal{I}$ for all $Y \subseteq X$, we have $\beta_\mu^a(Y) \equiv \mu\left(a^{-1}(Y)\right)$ and so in this case $\beta_\mu^a(\cdot)$ is a probability measure. More generally, for any arbitrary act a, $\beta_\mu^a(\cdot)$ is a *(simple) belief function* (that is, a special type of non-additive measure with finite support).

Definition 5 (Belief Function) *A simple belief function $\beta\colon 2^X \to [0\,,1]$ is a totally monotone normalized capacity with finite support. That is, $\beta(\emptyset) = 0$, $\beta(Y) = 1$, for some $Y \subset X$ with $|Y| < \infty$, $\beta(Z) \leqslant \beta(Z')$ whenever $Z \subset Z'$, and for any $k \geqslant 2$, and any collection of sets of outcomes $\{Y_1\,, \ldots\,, Y_k\}$ each $Y_k \subset X$,*

$$\beta\left(\bigcup_{j=1}^{k} Y_j\right) \geqslant \sum_{I \subseteq \{1,\ldots,k\}, I \neq \emptyset} (-1)^{|I|+1}\, \beta\left(\bigcap_{i \in I} Y_i\right).$$

Let Bel denote the set of belief functions (with finite support).[11]

In the sequel it will at times be useful to identify a belief function β with its associated *mass* function m.[12]

Definition 6 (Mass Function) *Fix a simple belief function $\beta \in Bel$ with support $Y \subset X$. The mass function m associated with β can be inductively defined on all the subsets of Y by setting:*

(i) $m(\emptyset) := 0$*;*

(ii) $m(\{x\}) := \beta(\{x\})$ *for each* $x \in Y$*; and,*

(iii) $m(Z) := \beta(Z) - \sum_{Z' \subset Z} m(Z')$*, for each* $Z \subseteq Y$*,* $|Z| > 1$*.*

[11] That $\beta_\mu^a(\cdot)$ satisfies all the properties required for a belief function is an immediate corollary of Proposition 3.1 in Fagin and Halpern (1991).

[12] The mass function is also known as the Möbius inverse.

Notice, since by construction we have $\beta(Z) = \sum_{Z' \subseteq Z} m(Z')$ for each $Z \subseteq Y$, the association between belief and mass functions is indeed one-to-one.

We shall denote by m^a_μ the mass function associated with the belief function β^a_μ to which the prior μ has mapped the act a. From its construction, it is natural to interpret $m^a_\mu(Y)$ as the weight of evidence the prior μ places on the act a leading to an outcome in the set Y occurring that has not already been assigned to some proper subset of Y.[13]

A consequence of the set $\mathcal{I}$ being closed under intersection is that the following Martingale property holds for the belief functions and hence for their associated mass functions.

Proposition 2 *Fix a prior μ. For any act $a \in A$ and any information event $R \in \mathcal{I}$:*

$$\beta^a_\mu = \mu(R)\beta^a_{\mu_R} + \mu(\Omega\backslash R)\beta^a_{\mu_{\Omega\backslash R}}\,; \text{ and,}$$
$$m^a_\mu = \mu(R)m^a_{\mu_R} + \mu(\Omega\backslash R)m^a_{\mu_{\Omega\backslash R}}\,.$$

Extending Machina and Schmeidler's (1992) concept of probabilistic sophisticated preferences to our setting in which not all events are deemed measurable by the DM, we require that her sole-option preferences admit a representation that along with her prior can be characterized by a suitably monotonic and continuous real-valued function W defined over belief functions. The continuity property we impose is mixture continuity, the analogue for belief functions of the one that Machina and Schmeidler impose on the corresponding function in their representation that is defined on probability distributions.

Definition 7 (Mixture Continuity) *A function $W\colon Bel \to \mathbb{R}$ is* mixture continuous *if for any three belief functions β, β' and β'' in $\Delta(\mathcal{X})$, the sets $\{\alpha \in [0,1]\colon W(\alpha\beta + (1-\alpha)\beta'') \geqslant W(\beta')\}$ and $\{\alpha \in [0,1]\colon W(\alpha\beta + (1-\alpha)\beta'') \leqslant W(\beta')\}$ are closed.*

The monotonicity property we impose relies on the following notion of outcome-set dominance.

Definition 8 (Outcome-Set Dominance) *We say an outcome set $Y = \{y_1, \dots, y_k\}$, with $y_i \succsim y_{i+1}$ for all $i = 1, \dots, k-1$,* dominates *outcome set $Z = \{z_1, \dots, z_n\}$, with*

[13] Hence, for any act a that is measurable with respect to μ, it follows that $m^a_\mu(Y) = 0$ for every $Y \subseteq a(\Omega)$ with $|Y| > 1$.

$z_j \succsim z_{j+1}$ *for all* $j = 1, \ldots, n-1$, *whenever either* : $k \leqslant n$ *and* $y_i \succ z_i$ *for each* $i = 1, \ldots, k$; *or,* $k > n$ *and* $y_{j+(k-n)} \succ z_j$ *for each* $j = 1, \ldots, n$.

To formally define our monotonicity property, we introduce some additional notation. For each non-empty $Y \subset X$, with $|Y| < \infty$, let β^Y denote the (degenerate) belief function that assigns 1 to any outcome set that is a superset of Y (and hence, assigns zero to any outcome set Z for which $Y \backslash Z \neq \emptyset$). That is, $\beta^Y(Z) := 1$ if $Z \supseteq Y$ and $\beta^Y(Z) := 0$ otherwise.

Definition 9 (Outcome-Set Monotonicity) *A function* $W : Bel \to \mathbb{R}$ *is* outcome-set monotonic *if for any belief function* $\beta \in B$, *and any pair of finite outcome-sets* Y *and* Z: Y *dominates* Z *implies*

$$W(\alpha\beta^Y + (1-\alpha)\beta) > W(\alpha\beta^Z + (1-\alpha)\beta), \text{ for all } \alpha \in (0, 1] .$$

To see the connection between our outcome-set monotonicity and Machina and Schmeidler's (1992) property of "monotonicity with respect to stochastic dominance" we define $\widehat{W}$, a function that maps mass functions to the reals, by setting $\widehat{W}(m) := W(\beta)$, where β is the belief function associated with the mass function m. For each non-empty $Y \subset X$, with $|Y| < \infty$, let m^Y denote the (degenerate) mass function that assigns 1 to Y (and hence, assigns zero to any outcome set Z not equal to Y). Notice that the degenerate belief function β^Y is associated with the degenerate mass function m^Y. With this identification in mind, it is straightforward to establish that W satisfies outcome-set monotonicity if and only if $\widehat{W}$ satisfies the following "degenerate independence" property: for any mass function m and any pair of finite subsets of X, Y and Z: Y dominates Z implies $\widehat{W}(\alpha m^Y + (1-\alpha)m) > \widehat{W}(\alpha m^Z + (1-\alpha)m)$, for all $\alpha \in (0, 1]$.

To interpret this, consider a pair of acts that are mapped by the prior μ to the respective belief functions $\alpha\beta^Y + (1-\alpha)\beta$ and $\alpha\beta^Z + (1-\alpha)\beta$. If m is the mass function associated with the belief function β, then the mass functions associated with these two acts are $\alpha m^Y + (1-\alpha)m$ and $\alpha m^Z + (1-\alpha)m$, respectively. Notice these two mass functions are identical except one has a weight α assigned to the outcome set Y while the other has that weight assigned to the outcome set Z. If Y dominates Z, then the first mass function can be seen to '(first-order) stochastically dominate' the second. Indeed if the two acts in question were both measurable with respect to μ then m would be a mass function that assigns positive weight only to some finite collection of singleton sets and we would have $Y = \{y\}$ and $Z = \{z\}$ for some pair of outcomes for which $y \succ z$. Furthermore, by a standard induction argument it follows the restriction of W to probability functions is *monotonic with respect to stochastic dominance* just as it is for Machina and Schmeidler's (1992) *probabistically sophisticated non-expected utility maximizer*.

We now have assembled all the elements required for a formal definition of the class of preferences that admit a representation characterized by a pair (μ, W), where μ is a prior and W is a mixture-continuous and outcome-set monotonic real-valued function (defined on belief functions).

Definition 10 (Generalized Expected Uncertain Utility Maximizers) *The sole-option preferences* $\succsim$ *belong to the class of* generalized expected uncertain utility maximizers *if there exists a prior* μ *and a mixture continuous and outcome-set monotonic real-valued function* W, *such that for every pair of acts* a *and* a' *in* A:

$$a \succsim a' \text{ if and only if } W(\beta_\mu^a) \geqslant W(\beta_\mu^{a'})\,.$$

Returning to our example from the introduction of DM's investment problem as described by the decision tree $\widehat{\tau}$, suppose DM's prior assigns a probability of 0.11 to the Republican winning the election (that is, $\mu(\sigma^{-1}(e)) = 0.11$) and hence assigns $\mu(\sigma^{-1}(d)) = 0.89$ to the Democrat winning. In addition, suppose investing in the wind farm yields DM a profit of 1 (billion dollars) irrespective of who wins the election. Alternatively, by investing in the small modular reactor DM will face, if the Republican wins, a gamble that yields a profit of either 5 with probability 10/11 or 0 with probability 1/11. Should the Democrat win, however, all DM can ascertain is that her profit from investing in the small modular reactor will be either 5 or 0. That is, given her perception of the nature of the uncertainty she faces under these circumstances, she is incapable of providing any more precise bounds on what the probability of either of these outcomes will be. The following table specifies the unconditional and conditional belief functions corresponding to DM's four possible behaviors in $\widehat{\tau}$.

Behavior $\left(b_e^{\widehat{\tau}}, b_d^{\widehat{\tau}}\right)$	Belief function if $\widehat{s}=e$	if $\widehat{s}=d$	unconditional
(wf, wf)	$\beta^{\{1\}}$	$\beta^{\{1\}}$	$\beta^{\{1\}}$
(wf, smr)	$\beta^{\{1\}}$	$\beta^{\{0,5\}}$	$\frac{11}{100}\beta^{\{1\}}+\frac{89}{100}\beta^{\{0,5\}}$
(smr, wf)	$\frac{1}{11}\beta^{\{0\}}+\frac{10}{11}\beta^{\{5\}}$	$\beta^{\{1\}}$	$\frac{1}{100}\beta^{\{0\}}+\frac{89}{100}\beta^{\{1\}}+\frac{1}{10}\beta^{\{5\}}$
(smr, smr)	$\frac{1}{11}\beta^{\{0\}}+\frac{10}{11}\beta^{\{5\}}$	$\beta^{\{0,5\}}$	$\frac{1}{100}\beta^{\{0\}}+\frac{89}{100}\beta^{\{0,5\}}+\frac{1}{10}\beta^{\{5\}}$

Recall, since her sole-option preferences yield the ranking

$$wf \succ smr_{\sigma^{-1}(e)}wf \succ smr \succ wf_{\sigma^{-1}(e)}smr\,,$$

it follows that

$$\begin{aligned} W\left(\beta^{\{1\}}\right) &> W\left(\frac{1}{100}\beta^{\{0\}}+\frac{89}{100}\beta^{\{1\}}+\frac{1}{10}\beta^{\{5\}}\right) \\ &> W\left(\frac{1}{100}\beta^{\{0\}}+\frac{89}{100}\beta^{\{0,5\}}+\frac{1}{10}\beta^{\{5\}}\right) > W\left(\frac{11}{100}\beta^{\{1\}}+\frac{89}{100}\beta^{\{0,5\}}\right) \end{aligned} \quad (1)$$

One can readily verify from these inequalities and the corresponding definitions of the conditional preferences obtained by applying resolute conditioning, that the ex ante optimal behavior $(wf\,, wf)$ is the only consistent plan of action. However, since the first inequality implies $wf \succ_e^{wf} smr$ while the third entails $smr \succ_e^{smr} wf$ we also see that the conditional preferences are not consequentialist. Indeed from (1) we see that the first inequality may be reexpressed as

$$W\left(\frac{11}{100}\beta^{\{1\}}+\frac{89}{100}\beta^{\{1\}}\right) > W\left(\frac{11}{100}\left(\frac{1}{11}\beta^{\{0\}}+\frac{1}{11}\beta^{\{5\}}\right)+\frac{89}{100}\beta^{\{1\}}\right)\,,$$

while the third can be reexpressed as

$$W\left(\frac{11}{100}\left(\frac{1}{11}\beta^{\{0\}}+\frac{1}{11}\beta^{\{5\}}\right)+\frac{89}{100}\beta^{\{0,5\}}\right) > W\left(\frac{11}{100}\beta^{\{1\}}+\frac{89}{100}\beta^{\{0,5\}}\right)\,.$$

Hence we can "explain" DM's sole-option preference ranking over the four possible courses of action in the tree $\widehat{\tau}$ as a manifestation of the *certainty effect* that has been offered as one explanation for the modal choice pattern seen in the Allais common-consequence paradox.

6 GEUU Maximization and Decomposability

Given a DM's sole-option preferences belongs to the class of GEUU maximizers and she engages in resolute conditioning (that is, Axiom 1 holds), for each information event $R \in \mathcal{I}$ and each baseline act $a \in A$, we see that $a' \succsim_R^a a''$ if and only if

$$W(\mu(R)\beta^{a'}_{\mu_R} + \mu(\Omega \backslash R)\, \beta^{a}_{\mu_{\Omega\backslash R}}) \geqslant W(\mu(R)\beta^{a''}_{\mu_R} + \mu(\Omega \backslash R)\, \beta^{a}_{\mu_{\Omega\backslash R}})\,. \tag{2}$$

Furthermore, from Theorem 1 we know for every possible behavior of the DM that constitutes a consistent plan to be ex ante optimal requires her sole-option preferences satisfy Axiom 3. This means for any four belief functions β, β', $\widehat{\beta}$, and $\widehat{\beta}'$ in Bel and any λ in $(0\,,1)$, we require

$$\begin{aligned} & W\left(\lambda\beta + (1-\lambda)\,\beta'\right) \geqslant \max\left\{\left(\lambda\widehat{\beta} + (1-\lambda)\,\beta'\right)\,, W\left(\lambda\beta + (1-\lambda)\,\widehat{\beta}'\right)\right\} \\ \Longrightarrow\ & W\left(\lambda\beta + (1-\lambda)\,\beta'\right) \geqslant W\left(\lambda\widehat{\beta} + (1-\lambda)\,\widehat{\beta}'\right) \end{aligned}$$

A property of W that is sufficient (but, we hasten to add, *not necessary*) to ensure this property holds, is linear additivity, namely,

$$W\left(\lambda\beta + (1-\lambda)\,\beta'\right) = \lambda W(\beta) + (1-\lambda)\,W\left(\beta'\right)\,. \tag{3}$$

For the following three subclasses of GEUU maximizers, the function W associated with the GEUU representation of their sole-option preferences is linearly additive.

Expected Uncertain Utility Maximization

In addition to a prior, the representation of Gul and Pesendorfer's (2014) EUU maximizer is characterized by an *interval utility* that corresponds to the DM's evaluation of a *set* of outcomes. It comprises a Bernoulli utility $v\colon X \to \mathbb{R}$ and a normalized (interval) aggregator $V\colon \{(u\,,u') \in v(X) \times v(X)\colon u \leqslant u'\} \to \mathbb{R}$ that is continuous and monotonic in the sense that $V(u\,,u') > V(\widehat{u}\,,\widehat{u}')$ whenever $u > \widehat{u}$ and $u' > \widehat{u}'$ and satisfies $V(u\,,u) = u$ for all $u \in V(X)$. And W takes the form

$$W(\beta) = \sum_{Y \subset X} V\left(\min_{x\in Y} v(x)\,, \max_{y\in Y} v(y)\right) m(Y)\,, \tag{4}$$

Hurwicz Expected Utility Maximization

A special case of EUU Maximization explored in detail in Gul and Pesendorfer (2015) is one in which the interval utility is additively separable. That is, $V(u\,,u') = \alpha u + (1-\alpha)u'$ for some α in $[0\,,1]$, hence

$$W(\beta) = \sum_{Y \subset X} \left(\alpha \min_{x\in Y} v(x) + (1-\alpha) \max_{y\in Y} v(y)\right) m(Y)\,, \tag{5}$$

Quasi-Arithmetic Mean Uncertain Utility Maximization

Eichberger and Pasichnichenko (2021) introduce and axiomatize a representation over belief functions in which the evaluation of a set of outcomes is performed by taking a *quasi-arithmetic* mean of the Bernoulli utilities of the outcomes in that set. So, in addition to a Bernoulli utility $v\colon X \to \mathbb{R}$ it requires a second-order utility $\phi\colon v(X) \to \mathbb{R}$ that is continuous and monotonic, in the sense that $\phi(w) > \phi(w')$ whenever $w > w'$, such that

$$W(\beta) = \sum_{Y \subset X} \phi^{-1}\left(\frac{1}{|Y|} \sum_{x\in Y} \phi\left(v(x)\right)\right) m(Y)\,. \tag{6}$$

Since we established in section 3 our DM's ranking of the four acts in $B(\widehat{\tau})$ violated **P2**, it follows that her sole-option preferences cannot come from any member of the class of GEUU maximizers for which W is linearly additive. However, the extension of the well-known *betweenness* property of Chew (1983) and Dekel (1986) to belief functions is both a necessary and sufficient for sole-option preferences that admit a GEUU representation to satisfy Axiom 3 (decomposability).

Definition 11 (Betweenness) *The function* $W\colon Bel \to \mathbb{R}$ *exhibits the* betweenness property*, if for all pairs of belief functions* β *and* β' *in* Bel*,* $W(\beta) \geqslant W(\beta') \implies W(\beta) \geqslant W\left(\lambda\beta + (1-\lambda)\beta'\right) \geqslant W(\beta')$*, for all* λ *in* $(0\,,1)$.

Proposition 3 *Suppose a DM's sole-option preferences* $\succsim$ *admit a GEUU representation characterized by a prior* μ *and mixture-continuous function* $W\colon Bel \to \mathbb{R}$*.*

Then the following are equivalent.

1. $\succsim$ *satisfies Axiom 3.*

2. W *exhibits the betweenness property.*

This proposition follows as a corollary to a representation result from a companion paper in which we established that Axiom 3 holds if and only if the (sole-option) preferences belong to the following subclass of GEUU maximizers.[14]

Implicit Linear Uncertain Utility Maximization

In addition to a prior, an ILUU representation of the sole-option preferences is characterized by a balanced outcome-set utility that represents the DM's evaluation of a *set* of outcomes *relative* to a balancing utility. It comprises a Bernoulli utility $v\colon X \to \mathbb{R}$ and a normalized (balanced) utility

$$U\colon \left\{(Y\,,u) : Y \subset X\,, |Y| < \infty\,, u \in \left[\inf_{x\in X} v(x)\,, \sup_{y\in X} v(y)\right]\right\} \to \mathbb{R}\,,$$

that satisfies the following four properties:

(*i*) *Balance Point Normalization*: for any $x \in X$: $U(\{x\}\,, v(x)) = 0\,.$

(*ii*) *Balance Point Monotonicity*: U is decreasing and continuous in its second argument.

(*iii*) *Outcome-Set Monotonicity:* for any u in the interval $\big(\inf_{x\in X} v(x)\,, \sup_{y\in X} v(y)\big)$, $U(Y\,,u) > U(Z\,,u)$, whenever Y dominates $Z\,.$

(*iv*) *Outcome-Set Continuity:* for any finite $Y \subset X$ and any sequence of outcome-sets $Z^n \subset X$ with $|Z^n| = |Y|$ for all n, if Z^n converges pointwise to Y, then $U(Z^n, u) \to U(Y, u)$ for all $u\,.$

For any belief function $\beta \in Bel$ with finite support $Y \subset X$, $W(\beta)$ is implicitly defined as the (unique) solution to

$$\sum_{Z\subseteq Y} U\left(Z\,, W(\beta)\right) m\left(Z\right) = 0\,, \tag{7}$$

[14] See Theorem 1 in Grant et al. (2022).

The explicit specifications of W in expressions (4) − (6) are all special cases of (7) in which $U(Y, v)$ equals $V(Y) - v$, for some outcome-set monotonic function V. This in turn implies W is linearly additive. The next example can be viewed as a parsimonious parameterization that cannot be expressed explicitly as a linearly additive function.

Disappointment Averse Hurwicz Expected Utility Maximization

Consider a special case of ILUU in which the balanced outcome-set utility is characterized by a Bernoulli utility $v\colon X \to \mathbb{R}$ and two parameters $\alpha \in [0, 1]$ and $\gamma > -1$ and takes the form

$$U(Y, u) = \begin{cases} \alpha \min_{y \in Y} v(y) + (1-\alpha) \max_{z \in Y} v(z) - u & \text{if } \alpha \min_{y \in Y} v(y) + (1-\alpha) \max_{z \in Y} v(z) \leqslant u \\ \frac{\alpha \min_{y \in Y} v(y) + (1-\alpha) \max_{z \in Y} v(z) - u}{1+\gamma} & \text{if } \alpha \min_{y \in Y} v(y) + (1-\alpha) \max_{z \in Y} v(z) > u \end{cases} \tag{8}$$

The restriction of the sole-option preferences to acts that are measurable with respect to $\mathcal{I}$ may be seen as conforming to Gul's (1991) risk preference model of *disappointment aversion.* Gul highlights a key feature of this model is the representation is only one parameter richer than expected utility. The analogous parameter here is γ (> -1). Indeed we see from expression (8), Hurzwicz Expected Utility corresponds to the case in which $\gamma = 0$.

The property of disappointment aversion that Gul shows is both necessary and sufficient to generate Allais style choice patterns requires $\gamma > 0$.[15] It is precisely this property that can accommodate the ranking by the DM's sole-option preferences of the acts in $B(\widehat{\tau})$. To see this, without loss of generality, we set $v(0) := 0$ and $v(5) := 1$. Furthermore, in order to make the calculations more straightforward let us also assume our DAHEU maximizing DM is maximally ambiguity averse, and so we set $\alpha := 1$. Our task is to find what values of $v(1)$ and $\gamma > 0$ allow the inequalities in (1) to hold. The table below lists the implicit characterization of the values of W for each of the four possible behaviors of DM in the tree $\widehat{\tau}$. For a fixed $\gamma > 0$, it readily

[15] We refer to a DM with $\gamma < 0$ as elation seeking.

follows from the four equations in this table that the inequalities in (1) hold if $v(1)$ lies in the open interval $\left(\frac{10}{11+\gamma}, \frac{89\gamma+100}{99\gamma+110}\right)$.

Characterizing $W(\cdot)$ for the four behaviors available to DM in the tree $B(\widehat{\tau})$.

Behavior $\left(b_e^{\widehat{\tau}}, b_d^{\widehat{\tau}}\right)$	$W\left(\beta_\mu^{\left(b_e^{\widehat{\tau}}, b_d^{\widehat{\tau}}\right)}\right)$ implicitly defined as solution to:
(wf, wf)	$v(1) - W\left(\beta_\mu^{(wf,wf)}\right) = 0$
(wf, smr)	$\frac{\left(v(1)-W\left(\beta_\mu^{(wf,smr)}\right)\right)}{1+\gamma}0.11 + \left(0 - W\left(\beta_\mu^{(wf,smr)}\right)\right)0.89 = 0$
(smr, wf)	$\left(0 - W\left(\beta_\mu^{(smr,wf)}\right)\right)0.01 + \frac{\left(v(1)-W\left(\beta_\mu^{(smr,wf)}\right)\right)}{1+\gamma}0.89 + \frac{\left(1-W\left(\beta_\mu^{(smr,wf)}\right)\right)}{1+\gamma}0.1 = 0$
(smr, smr)	$\left(0 - W\left(\beta_\mu^{(smr,smr)}\right)\right)0.90 + \frac{\left(1-W\left(\beta_\mu^{(smr,smr)}\right)\right)}{1+\gamma}0.1 = 0$

7 Concluding Comments

In a setup involving signal-contingent choices over menus of simple acts, we have shown that decomposability is the weakening of Savage's P2 that is not only still sufficient but also necessary for plans that are consistent to be ex ante optimal. The result is comparable to the characterization in Halevy (2004) of the absence of specualtive trades in terms of decomposability.

So decomposability allows for exactly that type of non-consequetialism of conditional choice, under which the conditional optimality of a plan, á la Nash, still guarantees global optimality. This observation leads to a straightforward verification procedure for the optimality of a plan, reminiscent of the considerations in Gul and Lantto (1990) that led those authors to employ the term 'dynamic programming solvability'. Indeed, echoing Grant et al. (2000), one can view decomposability as the analogue of dynamic programming solvability for the setting of subjective uncertainty where probabilities are not exogeneously specified.

References

Chew, S. H. (1983). A generalization of the quasilinear mean with applications to meaurement of income inequality and decision theory resolving the Allais paradox. *Econometrica 51*(4), 1065–1092.

Dekel, E. (1986). An axiomatic characterization of preferences under uncertainty: weakening the independence axiom. *Journal of Economic Theory 40*, 304–318.

Eichberger, J. and I. Pasichnichenko (2021, December). Decision making with partial information. *Journal of Economic Theory 198*.

Ellsberg, D. (1961). Risk, ambiguity, and the Savage axioms. *Quarterly Journal of Economics 75*, 643–669.

Fagin, R. and J. Halpern (1991). Uncertainty, belief, and probability. *Computational Intelligence 7*, 160–173.

Grant, S., A. Kajii, and B. Polak (1998). Intrinsic preference for information. *Journal of Economic Theory 83*, 233–259.

Grant, S., A. Kajii, and B. Polak (2000). Decomposable Choice under Uncertainty. *Journal of Economic Theory 92*, 167–197.

Grant, S., B. Roorda, and J. Yang (2022, June). Balanced Set Utility – EUU theory with set outcomes and betweenness.

Gul, F. (1991). A Theory of Disappointment Aversion. *Econometrica 59*(3), 667–686.

Gul, F. and O. Lantto (1990). Betweenness satisfying preferences and dynamic choice. *Journal of Economic Theory 52*(1), 162–177.

Gul, F. and W. Pesendorfer (2014). Expected uncertain utility. *Econometrica 82*(1), 1–39.

Gul, F. and W. Pesendorfer (2015). Hurwicz expected utility and subjective sources. *Journal of Economic Theory 159*, 465–488.

Halevy, Y. (2004). The possibility of speculative trade between dynamically consistent agents. *Games and Economic Behavior 46*, 189–198.

Johnsen, T. and J. Donaldson (1985). The structure of intertemporal preferences under uncertainty and time consistent plans. *Econometrica 53*, 1451–1458.

Kahneman, D. and A. Tversky (1981). The framing of decisions and the psychology of choice. *Science 211*, 453–458.

Karni, E. and D. Schmeidler (1991). Atemporal dynamic consistency and expected utility theory. *Journal of Economic Theory 54*, 401–408.

Kopylov, I. (2007). Subjective probabilities on "small" domains. *Journal of Economic Theory 133*(1), 236–265.

Machina, M. J. (1989). Dynamic consistency and non-expected utility models of choice under uncertainty. *Journal of Economic Literature 27*, 1622–1668.

Machina, M. J. and D. Schmeidler (1992). A More Robust Definition of Subjective Probability. *Econometrica 60*(4), 745–780.

McClennen, E. (1990). *Rationality and Dynamic Choice: Foundational Explorations*. Cambridge, UK: Cambridge Univeristy Press.

Nesterov, Y. (2012). Efficiency of coordinate descent methods on huge-scale optimization problems. *SIAM Journal on Optimization 22*(2), 341–362.

Savage, L. J. (1954). *The Foundations of Statistics*. John Wiley & Sons.

Siniscalchi, M. (2011). Dynamic choice under ambiguity. *Theoretical Economics 6*, 379–421.

Strotz, R. (1955). Myopia and Inconsistency in Dynamic Utility Maximization. *Review of Economic Studies 23*(3), 165–180.

Wright, S. J. (2015). Coordinate descent algorithms. *Mathematical Programming 151*(1), 3–34.