Maxmin Weighted Expected Utility: A Simpler Characterization*

Joseph Y. Halpern
Cornell University
halpern@cs.cornell.edu

Samantha Leung
Cornell University
samlyy@cs.cornell.edu

August 7, 2015

Abstract

Chateauneuf and Faro [3] axiomatize a weighted version of maxmin expected utility over acts with nonnegative utilities, where weights are represented by a confidence function. We argue that their representation is only one of many possible, and we axiomatize a more natural form of maxmin weighted expected utility. We also provide stronger uniqueness results.

1 Introduction

Maxmin expected utility (MMEU), axiomatized by Gilboa and Schmeidler [6], is one of the best-studied alternatives to subjective expected utility (SEU) maximization [13]. Its compatibility with *ambiguity-averse* preferences makes it an attractive descriptive decision model, in light of experimental evidence (e.g., the Allais Paradox [1] and the Ellsberg Paradox [5]) showing that intuitive decisions may violate the ambiguity neutrality, or "independence", property implied by the SEU model. In the (multiple priors) MMEU decision model, there is a set of possible probability distributions over the statespace, each giving rise to a (potentially different) expected utility value for each object of choice. An MMEU decision maker chooses an option that maximizes the minimum of such expected utility values.

However, even MMEU may be too restrictive a model for representing reasonable decision-making. For example, Chateauneuf and Faro

*Work supported in part by NSF grants IIS-0812045, IIS-0911036, and CCF-1214844, by AFOSR grants FA9550-08-1-0438, FA9550-09-1-0266, and FA9550-12-1-0040, and by ARO grant W911NF-14-1-0017.

[3] (henceforth CF) point out that MMEU does not allow for "attraction for smoothing an uncertain act with the help of a positive constant act", a property that is intuitively reasonable and is demonstrated in Example 5.2.

To deal with this, CF consider a "weighted" version of maxmin expected utility [6]. Recall that in the MMEU model, beliefs are represented by a set of probability measures over the state space. The distributions that are in the set are viewed as the possible distributions over the states. However, sometimes it makes sense to treat some distributions as "more likely" than other distributions, rather than just separating the distributions into two groups ("possible" and "impossible"). CF provide a method of treating distributions differently, by assigning a *confidence value* to each distribution.

Others have independently studied similar models. Klibanoff et al. [9] propose a model of decision making that associates weights with probability measures, but makes decisions based on a "weighted" expected utility function. Maccheroni et al. [10] study a model of decision making where additive, instead of multiplicative, weights are associated with probability measures. Hayashi [8] considers a model of expected regret minimization where the regret associated with each state is taken to a positive power before the expectation is taken. In previous work [7] we have also considered associating multiplicative weights with probability measures in expected-regret-minimization. Others have also proposed and studied approaches of representing uncertainty that are similar to weighted probabilities (see, e.g. [4, 11, 14]).

In the CF model, a high confidence value on a probability measure can be interpreted as the probability measure being "significant" or "likely to be the correct distribution," while a low confidence value on a probability measure is interpreted as the probability measure being insignificant or unlikely to be the correct distribution. These confidence values are used to scale the expected utilities of the acts in a way that reflects the relative significance of each probability measure. Since larger weights should always magnify the influence of a distribution, one must restrict to either nonnegative or nonpositive utilities. CF choose to restrict to nonnegative utilities, and they multiply the expected utilities by the multiplicative inverse of the associated confidence value. The maxmin expected utility criterion is then used to compare utility acts based on these "weighted" expected utilities. In this paper, we use the term *weight* to refer to the final real number that we multiply the expected utilities by. In the CF model, the weight is obtained by taking the multiplicative inverse of the confidence value. Multiplying by the inverse ensures that probability measures with low confidence have a smaller effect, since they are less likely to give the minimum expected utility. This generalization of the maxmin expected utility decision rule allows for a "smoothing" effect. Instead of simply being in or out of the set of probability measures considered possible, probability measures now have finer weights associated with them.

However, CF also introduce a numerical confidence threshold $\alpha_0 > 0$; a probability measure is "discarded" (i.e., ignored) if its confidence value is below this threshold α_0. This threshold affects the resulting behavior of the decision model, as captured by the axioms characterizing the decision model. Having this threshold seems to us incompatible with the intuition behind weights. If a probability measure has low weight, we should perhaps take it less seriously than one with high weight, but there seems to be no good reason to ignore it altogether. Therefore, we define a simpler version of the decision rule where there is no threshold α_0. This simplified decision rule is characterized by removing one of the CF axioms.

Another problem with the CF approach is that of using the multiplicative inverse of the confidence value as the weight on the expected utilities. This choice seems rather arbitrary. Why not use the square of the inverse? We show that any monotonically decreasing transformation that maps $(0, 1]$ onto $\mathbb{R}^+$ (the nonnegative reals) satisfies the same axioms. Although all these transformations are characterized by the same axioms, different transformations may lead to quite different decisions.

It is not clear which transformation function is the "right" one. There is no compelling argument for using $\frac{1}{x}$ rather than, say, $\frac{1}{x^2}$. Our axiomatization leads to some important observations:

1. What is important is the composition $t \circ \phi$ of the transformation function t and the confidence function ϕ, not the confidence function itself nor the transformation function itself; it is the composition that determines the preferences.

2. Confidence values have no cardinal meaning: a confidence value of $\frac{1}{2}$ can have the same meaning as a confidence value of $\frac{1}{3}$ if the transformation t changes.

Moreover, as our results show, the confidence value and the transformation interact. In our earlier work on minimax weighted expected regret [7], we were able to get a strong uniqueness result in the context of regret by multiplying the probability measure by the weight. That is, instead of considering the set of probability measures and the associated weights separately, we consider what we called subprobability measures, which are probability measures "scaled" by a weight in $[0, 1]$. By looking at these subprobbility measures, we were able to find natural properties to ensure uniqueness of the representation. Here, we show that by multiplying the probability measure by the weight, we can get a uniqueness result analogous to that for regret.

With weighted regret, there is no need to apply a transformation to the confidence values. The weights are simply the confidence values. Equivalently, the identify function is a valid transformation for weighted regret. We show that for maxmin weighted expected utility, if we restrict to *nonpositive* utilities instead of nonnegative utilities, we can also take the transformation to be the identity function. That is, we can just multiply the expected utilities by a confidence value without applying any transformations. We then replace the axiom saying that there is a worst outcome with one saying that there is a best outcome. This results in essentially the same representation theorem.

The rest of this paper is organized as follows. Section 2 sets up some preliminary definitions. Section 3 presents the CF model and some of their results. Section 4 considers a generalization of the CF model. Section 5 presents a simpler model and provides a representation theorem. Proofs are collected in the appendices.

2 Formal Definitions

In this section we provide definitions that will be used to present the CF results, as well as to develop our new results. We restrict to what is known in the literature as the *Anscombe-Aumann* (AA) framework [2], where outcomes are restricted to lotteries. This framework is standard in the decision theory literature; axiomatic characterizations of SEU [2] and MMEU [6] have been obtained in the AA framework.

We assume that the state space S is associated with a sigma algebra, and we let $\Delta(S)$ denote the set of all probability distributions on S. Given a set X (which we view as consisting of *prizes* or *outcomes*), a *lottery* over X is just a probability distribution on X with finite support. Let $\underline{\Delta}(X)$ be the set of all lotteries. In the AA framework, the set of outcomes is $\underline{\Delta}(X)$. So now acts are functions from the state space S to $\underline{\Delta}(X)$. (Such acts are sometimes called *Anscombe-Aumann acts*.) We denote the set of all acts by $\mathcal{F}$. The technical advantage of considering such a set of outcomes is that we can consider convex combinations of acts. If f and g are acts, define the act $\alpha f + (1-\alpha)g$ to be the act that maps a state s to the lottery $\alpha f(s) + (1-\alpha)g(s)$.

Given a utility function U on prizes in X, the utility of a lottery $l \in \underline{\Delta}(X)$ is just the expected utility of the prizes obtained, that is,

$$u(l) = \sum_{\{x \in X \,:\, l(x) > 0\}} l(x)U(x).$$

This makes sense since $l(x)$ is the probability of getting prize x if lottery l is played. The expected utility of an act f with respect to a probability p on states is then just $u(f) = \int_S u(f(s))dp$, as usual.

3 CF Maxmin Expected Utility with Confidence Functions

The CF approach is formalized as follows. Let $\phi : \Delta(S) \to [0,1]$ be a confidence function on the probability measures, and let u be a utility function on lotteries over X with values in $\mathbb{R}^+$ (all instances of $\mathbb{R}^+$ in this paper include 0). Let $L_{\alpha_0}\phi$ denote the set $\{p \in \Delta(S) : \phi(p) \geq \alpha_0\}$ for $\alpha_0 \in (0,1]$.

Definition 3.1. *Define* $\succeq_\phi^{+,\alpha_0}$ *so that*

$$f \succeq_\phi^{+,\alpha_0} g \Leftrightarrow \min_{p \in L_{\alpha_0}\phi} \frac{1}{\phi(p)} \int_S u(f)dp \geq \min_{p \in L_{\alpha_0}\phi} \frac{1}{\phi(p)} \int_S u(f)dp.$$

The superscript + on $\succeq_\phi^{+,\alpha_0}$ indicates that the preference is defined for nonnegative utilities. Note that, according to Definition 3.1, a probability measure that has a confidence value (according to ϕ) lower than α_0 is simply discarded. The analogy to maxmin expected utility of Gilboa and Schmeidler [6] is that the probability measure is not in the belief set. Indeed, if $\alpha_0 = 1$, then the CF approach essentially reduces to maxmin expected utility.

CF call confidence functions satisfying the following properties *regular* fuzzy sets.*

Definition 3.2. *The set of regular* fuzzy sets consists of all mappings* $\phi : \Delta(S) \to [0,1]$ *satisfying the following properties:*

(a) ϕ *is* normal*:* $\{p \in \Delta(S) : \phi(p) = 1\} \neq \emptyset$.

(b) ϕ *is* weakly* upper semicontinuous*:* $\{p \in \Delta(S) : \phi(p) \geq \alpha\}$ *is weakly* closed for all* $\alpha \in [0,1]$.

(c) ϕ *is* quasi-concave*:*

$$\forall \beta \in [0,1](\phi(\beta p_1 + (1-\beta)p_2) \geq \min\{\phi(p_1), \phi(p_2)\}).$$

One role of regular* fuzzy sets in the CF representation is that the condition provides a canonical representation. That is, every preference order satisfying appropriate axioms can be represented by some utility function, some $\alpha_0 > 0$, and some regular* fuzzy ϕ. Moreover, there is a ϕ^* within the set of regular* fuzzy sets generating these preferences such that ϕ^* is maximal in the sense that for every probability measure p, ϕ^* assigns weakly larger confidence to p than every other regular* fuzzy set generating these preferences.

CF consider the following axioms. In the axioms, the acts f and g are viewed as being universally quantified; given an outcome $x \in X$, we write x^* to denote the constant act that maps all states to the outcome x.

Axiom 1.
a. (Transitivity)*:* $f \succeq g \succeq h \Rightarrow f \succeq h$.
b. (Completeness)*:* $f \succeq g$ *or* $g \succeq f$.
c. (Nontriviality)*:* $f \succ g$ *for some acts* f *and* g.

Axiom 2 (Monotonicity)**.** *If* $(f(s))^* \succeq (g(s))^*$ *for all* $s \in S$*, then* $f \succeq g$.

Axiom 3 (Continuity)**.** *For all* $f, g, h \in \mathcal{F}$*, the sets* $\{\alpha \in [0,1] : \alpha f + (1-\alpha)g \succeq h\}, \{\alpha \in [0,1] : h \succeq \alpha f + (1-\alpha)g\}$ *are closed.*

Axiom 4 (Worst Independence)**.** *There exists a* worst outcome $\underline{x} \in X$ *such that* $f \succeq \underline{x}^*$ *for every* $f \in \mathcal{F}$*. Moreover,*

$$f \sim g \Rightarrow \alpha f + (1-\alpha)\underline{x}^* \sim \alpha g + (1-\alpha)\underline{x}^*.$$

Axiom 4 is reminiscent of Gilboa and Schmeidler's [6] C-independence axiom of MMEU; C-independence is stronger in the sense that the independence property needs to hold not only for $\underline{x}^*$, but all other constant acts as well.

Axiom 5 (Independence on Constant Acts)**.**

$$\forall x, y, z \in X (x^* \sim y^* \Leftrightarrow \frac{1}{2}x^* + \frac{1}{2}z^* \sim \frac{1}{2}y^* + \frac{1}{2}z^*).$$

Axiom 5 is a weaker version of the more common independence axiom for constant acts, where instead of $\frac{1}{2}$ mixtures, all convex mixtures of the constant acts are allowed. CF chose to present this weaker axiom, since it was shown by Herstein and Milnor [**?**] that Axioms 1, 3 and 5 are sufficient to satisfy the premises of the von-Neumann-Morgenstern theorem, which says that there is an expected-utility representation for preferences over constant acts. While we could have used the more standard/stronger versions of the continuity and independence axioms, to make comparisons easier, we use the versions used by CF.

Axiom 6 (Ambiguity Aversion)**.**

$$f \sim g \Rightarrow pf + (1-p)g \succeq g.$$

Ambiguity aversion says that when there are two equally good alternatives, the decision maker prefers to hedge between these two alternatives. Ambiguity aversion is also sound for MMEU [6].

Axiom 7 (Bounded Attraction for Certainty)**.** *There exists* $\delta \geq 1$ *such that for all* $f \in \mathcal{F}$ *and* $x, y \in X$*:*

$$x^* \sim f \Rightarrow \frac{1}{2}x^* + \frac{1}{2}y^* \succeq \frac{1}{2}f + \frac{1}{2}\left(\frac{1}{\delta}y^* + \left(1 - \frac{1}{\delta}\right)\underline{x}^*\right).$$

As CF point out, Axiom 6 implies that if an agent is indifferent between an act f and a constant act x^*, then she could strictly the convex combination of f with a constant act y^* to the combination of x^* and y^*. In particular, if we let $y^* = x^*$, then Axiom 6 implies that $pf + (1-p)y^* \succeq x^* = px^* + (1-p)y^*$ for all $p \in [0,1]$. CF explain that Axiom 7 imposes a bound on the affinity for smoothing out an uncertain act with a constant act. Continuing with our example and letting $\underline{x}^* = 0^*$ (assuming that outcomes are numbers), Axiom 7 implies that $\frac{1}{2}x^* + \frac{1}{2}y^* \succeq \frac{1}{2}f + \frac{1}{2\delta}y^*$ for some fixed δ specified by Axiom 7. The fact that there exists a $\delta > 1$ such that $\frac{1}{2}x^* + \frac{1}{2}y^* \succeq \frac{1}{2}f + \frac{1}{2\delta}y^*$ follows from monotonicity. The power of Axiom 7 comes from the fact that there is a single $\delta \geq 1$ such that this preference holds for all $x, y \in X$, and $f \in \mathcal{F}$.

The Bounded Attraction for Certainty axiom in the CF representation captures the lower bound α_0 in the model. Recall that if the confidence value of a probability measure is less than α_0, then that measure is considered "impossible", or ignored. CF show that the δ in the Bounded Attraction for Certainty axiom can be taken to be $\frac{1}{\alpha_0}$ in the representation. δ is roughly interpreted as an upper bound on how much the mixing of a constant act to an act can make the act more preferable. We essentially take $\alpha_0 = 0$; all probability measures into account, regardless of their weight, as long as the weight is positive. Since weighted regret already says that regret due to probability measures with low confidence is not taken seriously, there seems to be no reason to ignore probability measures of low confidence altgether. In any case, since we take $\alpha_0 = 0$, we would expect decision rule to satisfy an unbounded version of attraction for certainty. Our representation theorem shows that such an axiom is not needed to characterize maxmin weighted expected utility.

CF prove the following representation theorem:

Theorem 3.3 (CF representation theorem [3])**.** *A binary relation $\succeq$ on $\mathcal{F}$ satisfies Axioms 1–7 if and only if there exists a unique non-constant function $u : X \to \mathbb{R}^+$ such that $u_{x_*} = 0$, unique up to positive linear transformations, a minimal confidence level $\alpha_0 \in (0,1]$, and a regular* fuzzy set $\phi : \Delta(S) \to [0,1]$ such that $\succeq = \succeq_\phi^{+,\alpha_0}$.*

Note that although CF guarantee the existence of a representation with a regular* fuzzy set, the confidence function does not necessarily need to be regular* fuzzy in order to satisfy Axioms 1–7. For example, if there are two states, s_1 and s_2, p_i is the point-mass on state s_i for $i \in \{1,2\}$, $\phi(p_1) = \phi(p_2) = 1$, and $\phi(p) = 0$ for all other probability measures p, then ϕ is not a regular* fuzzy set, since it is not quasi-concave. Nevertheless, $\succeq_\phi^{+,\frac{1}{2}}$ is determined by maxmin expected utility and thus must satisfy Axioms 1–7, because Axioms 1–7 are strictly weaker than the axioms for maxmin expected utility [6].

4 t-Maxmin Weighted Expected Utility

In this section we consider a generalization of the CF approach, which we call the *t-maxmin weighted decision rule.* The t-maxmin weighted rule applies a monotonically decreasing transformation function t to the confidence values, and then uses the maxmin criterion on expected utilities multiplied by the transformed confidence values. The CF decision rule is the special case of the t-weighted maxmin decision rule, where $t(x) = \frac{1}{x}$.

Let $\phi : \Delta(S) \to [0, 1]$ be a confidence function, let $t : (0, 1] \to \mathbb{R}^+$ be a transformation function, and let u be a nonnegative utility function.

Definition 4.1 (t-maxmin weighted expected utility)**.** *Define* $\succeq_{t,\phi}^{+,\alpha_0}$ *so that*

$$f \succeq_{t,\phi}^{+,\alpha_0} g \Leftrightarrow \min_{p \in L_{\alpha_0}\phi} t(\phi(p)) \int_S u(g)dp \geq \min_{p \in L_{\alpha_0}\phi} t(\phi(p)) \int_S u(f)dp.$$

The threshold value α_0 affects the preferences $\succeq_{\phi}^{+,\alpha_0}$ only if it is larger than the smallest confidence value. That is, let $\alpha_0^*(\phi) = \max\{\alpha_0, \inf_{p\in\Delta(S)} \phi(p)\}$. It is easy to see that, for all $0 < \alpha \leq \alpha_0^*(\phi)$, we have $\succeq_{\phi}^{+,\alpha} = \succeq_{\phi}^{+,\alpha_0^*(\phi)}$.

Theorem 4.2 shows that it is not necessary to use the transformation $t(x) = \frac{1}{x}$ to map confidence values into weights with which expected utilities are multiplied. Other functions, such as $t(x) = \frac{1}{x^2}$, represent the same class of preference orders. However, there are some constraints on the allowed transformation functions t, since we need to "simulate" $\frac{1}{\phi(p)}$ with $t(\phi'(p))$. In addition to being strictly decreasing (a property of $t(x) = \frac{1}{x}$), the condition that there exists some $\beta > 0$ such that $[\beta, \beta/\alpha_0^*(\phi)] \subseteq range(t)$ guarantees that we can "simulate" $\frac{1}{\phi(p)}$ with $t(\phi'(p))$ for some ϕ' and α_0'. Continuity guarantees that we can find a preimage $\phi'(p)$ for every value in the range of t.

Theorem 4.2. *For all measurable spaces* (S, Σ)*, consequences* X*, nonnegative utility functions* u*, confidence functions* $\phi : \Delta(S) \to [0, 1]$*, thresholds* $\alpha_0 > 0$*, and strictly decreasing, continuous transformation functions* $t : (0, 1] \to \mathbb{R}^+$ *such that there exists some* $\beta > 0$ *such that* $[\beta, \beta/\alpha_0^*(\phi)] \subseteq range(t)$*, there exists* $\alpha_0' > 0$ *and* ϕ' *such that*

$$\succeq_{\phi}^{+,\alpha_0} = \succeq_{t,\phi'}^{+,\alpha_0'};$$

moreover, if ϕ *is regular** *and* $t(1) = \beta$*, then* ϕ' *is regular*.*

Theorem 4.2 highlights another perspective of the t-weighted maxmin expected utility representation. In addition to viewing $\phi : [0, 1]$ as a confidence function which is transformed and then applied to probability measures, we can also view $t(\phi(p))$ as a *weight* applied to the probability measure p. In this paper, we use the term *weight* to refer to a value in $\mathbb{R}^+$ with which the expected probability is multiplied, while the term *confidence* refers to a value in $[0, 1]$ in the sense used by Chateaneuf and Faro. In the theorem statement (and later in the paper), we take U^+ to denote a nonnegative utility function.

A corollary of Theorem 4.2 is a representation theorem for the CF axioms, that is, Axioms 1–7. Theorem 4.3 requires that $t(1) > 0$, since if $t(1) \leq 0$ and the confidence function is normal then the preferences will be trivial. Theorem 4.3 provides a stronger uniqueness result than Theorem 3.3.

Theorem 4.3. *Let $t : (0, 1] \to \mathbb{R}^+$ be a continuous, strictly decreasing function with $t(1) > 0$ and $\lim_{x\to 0^+} t(x) > c$ for $c \in \mathbb{R}^+$. For all X, U^+, S, $\alpha_0 > 0$, and ϕ, if U^+ is nonconstant and $\alpha_0^*(\phi) \geq c$, then the preference order $\succeq_{t,\phi}^{+,\alpha_0}$ satisfies Axioms 1–7, with $\delta = \frac{c}{t(1)}$ in Axiom 7. Conversely, if the preference order $\succeq$ on the acts in $\mathcal{F}$ satisfies Axioms 1–7 with $t(1)\delta \leq c$ in Axiom 7, then there exists a nonnegative utility function U^+ on X, a threshold $\alpha_0 > 0$, and a confidence function $\phi : \Delta(S) \to [0, 1]$ such that ϕ is regular* fuzzy, $t \circ \phi$ has convex upper support, and $\succeq = \succeq_{t,\phi}^{+,\alpha_0}$. Moreover, U^+ is unique up to positive linear transformations, and if S is finite, there is a sense in which ϕ is unique (see Theorem 5.5).*

Proof. That $\succeq_{t,\phi}^{+,\alpha_0}$ satisfies Axioms 1–7 follows from Theorem 3.3 and Theorem 4.2, since $\succeq_{t,\phi}^{+,\alpha_0} = \succeq_{\phi'}^{+,\alpha_0'}$ for some α_0' and ϕ', and $\succeq_{\phi'}^{+,\alpha_0'}$ satisfies Axioms 1–7.

Proving the converse also involves Theorems 3.3 and 4.2. If a preference order satisfies Axioms 1–7, then by Theorem 3.3 there exists a CF representation. Moreover, the α_0 in the construction of the representation in CF's proof of Theorem 3.3 is equal to $\frac{1}{\delta}$, where δ is the number in Axiom 7. Also recall that $\alpha_0 \leq \alpha_0^*$. Therefore, if $\lim_{x\to 0^+} t(x) > t(1)\delta$ and $t(1) > 0$, then for $\beta = t(1)$, we have $[\beta, \beta/\alpha_0^*(\phi)] \subseteq [\beta, \beta\delta] \in range(t)$ over the domain $(0, 1]$. By Theorem 4.2, we can conclude that there exists a t-weighted maxmin expected utility representation.

The uniqueness claim follows from Theorem 5.5 below, which requires only Axioms 1–6. □

It is well known that for MMEU and regret, the preference order determined by a set P of probability measures is the same as that determined by the convex hull of P. Thus, to get uniqueness, Gilboa and Schmeidler [6] consider only convex sets of probability measures. In [7], we show that a set of sub-probability measures determine the same minimax weighted expected regret (MWER) preferences as its convex hull. Proposition 4.5 shows that the generalized probability measures behave in much the same way as the probability measures in MMEU and the sub-probability measures in MWER.

Given a set V of generalized probabilities, define the relation $\succeq_V$ by taking

$$f \succeq_V g \Leftrightarrow \inf_{p\in V} \int_S u(f)dp \geq \inf_{p\in V} \int_S u(g)dp.$$

It is not difficult to see that we can convert back and forth between the upper support of a weighting function and the weighting function itself. Therefore, we lose no information by looking at the upper support of a weighting function.

Proposition 4.4. $\succeq_{\overline{V}^{\alpha_0}_{t\circ\phi}} = \succeq^{+,\alpha_0}_{t,\phi}$.

Proof.

$$\begin{aligned} f \succeq_{\overline{V}^{\alpha_0}_{t\circ\phi}} g \text{ iff } & \inf_{p'\in\overline{V}^{\alpha_0}_{t\circ\phi}} \int_S u(f)dp' \geq \inf_{p'\in\overline{V}^{\alpha_0}_{t\circ\phi}} \int_S u(g)dp' \\ \text{iff } & \inf_{\{q:q=t(\phi(p))p,\phi(p)>\alpha_0\}} \int_S u(f)dq \geq \inf_{\{q:q=t(\phi(p))p,\phi(p)>\alpha_0\}} \int_S u(g)dq \\ \text{iff } & \inf_{\{p:\phi(p)>\alpha_0\}} t(\phi(p)) \int_S u(f)dp \geq \inf_{\{p:\phi(p)>\alpha_0\}} t(\phi(p)) \int_S u(f)dp \\ \text{iff } & f \succeq^{+,\alpha_0}_{t,\phi} g, \end{aligned}$$

if $\phi(p)$ is lower semi-continuous. □

Recall that, given a set V in a mixture space, $Conv(V) = \{\alpha x + (1-\alpha)y : x, y \in V, \alpha \in [0,1]\}$ is the convex hull of V.

Proposition 4.5. *If V, V' are sets of generalized probability measures and $Conv(V) = Conv(V')$, then $\succeq_V = \succeq_{V'}$.*

Proof. It suffices to show that V represents the same preferences as $Conv(V)$. Let V be a set of generalized probability measures. Given $\beta \in [0,1]$, $p_1, p_2 \in V$, and an act $f \in \mathcal{F}$, we have

$$\beta \int u(f)dp_1 + (1-\beta) \int u(f)dp_2 \geq \min\{\int u(f)dp_1, \int u(f)dp_2\}.$$

This means that $\beta p_1 + (1-\beta)p_2$ can be added to V without changing the preferences, as required. □

4.1 Impact of the threshold

In the following example, we examine how Axiom 7 qualitatively affects the weighted maxmin expected utility preferences.

Example 4.6. Suppose there are two states: $S = \{s_0, s_1\}$. Consider the confidence function ϕ defined by $\phi(p) = \sqrt{p(s_1)}$. Like CF, we let $t(x) = \frac{1}{x}$, and let $\alpha_0 > 0$ be a fixed threshold value. Let $\succeq_\phi^{+,\alpha_0}$ be resulting preference relation. Let f be an act such that $u(f(s_0)) = 0$ and $u(f(s_1)) = 1$. Let c^* be a constant act with utility $c > 0$. Then we have that

$$f \succeq_\phi^{+,\alpha_0} c^* \Leftrightarrow \inf_{\{p: \sqrt{p(s_1)} \geq \alpha_0\}} \sqrt{p(s_1)} \geq c.$$

This means that f is strictly preferred to all constant acts c^* with $c < \alpha_0$, but is considered strictly worse than all constant acts c^* with $c > \alpha_0$.

Now compare this to the preference order obtained by considering the same confidence function c and weight function t, but with no threshold on the confidence. Then we have that

$$f \succeq_\phi^{+} c^* \Leftrightarrow \inf_{p \in \Delta(S)} \sqrt{p(s_1)} \geq c.$$

Since $\min_{p \in \Delta(S)} \sqrt{p(s_1)} = 0$, this means that f is strictly worse than all constant acts c with $c > 0$. Clearly, imposing a threshold has a nontrivial impact on the preference order.

We can also show how CF's Axiom 7 is violated by $\succeq_\phi^{+}$. Suppose that the worst outcome in this example (i.e., $\underline{x}$) is 0. If there is no threshold (or, equivalently, if $\alpha_0 = 0$), then $f \sim 0^*$. Thus, Axiom 7 implies that, for some fixed $\epsilon > 0$, for all outcomes y, we have that $\frac{1}{2}y^* \succeq \frac{1}{2}f + \epsilon y^*$. However,

$$\begin{array}{lll} & \frac{1}{2}y^* & \succeq_\phi^{+,0} \frac{1}{2}f + \epsilon y^* \\ \text{iff} & \frac{y}{2} & \geq \inf_{p \in \Delta(S)} \left(\frac{1}{\sqrt{p(s_1)}} (p(s_1)(\frac{1}{2} + \epsilon y) + (1 - p(s_1))\epsilon y) \right) \\ & & = \inf_{p \in \Delta(S)} \left(\frac{\epsilon y}{\sqrt{p(s_1)}} + \frac{1}{2}\sqrt{p(s_1)} \right) \end{array}$$

It is easy to see that

$$\inf_{p \in \Delta(S)} \left(\frac{\epsilon y}{\sqrt{p(s_1)}} + \frac{1}{2}\sqrt{p(s_1)} \right) = \sqrt{2\epsilon y},$$

which means that for all $y < 8\epsilon$, we have that $\frac{1}{2}y \prec \frac{1}{2}f + \epsilon y$, contradicting Axiom 7.

5 Maxmin Weighted Expected Utility

5.1 Removing the threshold

As discussed in the previous section, it does not seem natural to discard probability measures if their confidence values do not meet some fixed threshold $\alpha_0 > 0$. We can naturally extend the definition of t-weighted maxmin expected utility to remove the threshold α_0.

Definition 5.1 (t-maxmin weighted expected utility without α_0)**.** *Define $\succeq^+_{t,\phi}$ so that*

$$f \succeq^+_{t,\phi} g \Leftrightarrow \inf_{\{p:\phi(p)>0\}} t(\phi(p)) \int_S u(g)dp \geq \inf_{\{p:\phi(p)>0\}} t(\phi(p)) \int_S u(f)dp.$$

Clearly $\succeq^{+,\alpha_0}_{t,\phi} = \succeq^+_{t,\phi'}$ where $\phi'(p) = \phi(p)$ if $\phi(p) \geq \alpha_0$ and $\phi'(p) = 0$ if $\phi(p) < \alpha_0$. Thus, $\succeq^+_{t,\phi}$ is at least as expressive as $\succeq^{+,\alpha_0}_{t,\phi}$.

If we consider CF's preference order $\succeq^+_\phi$ without a threshold α_0, then as Example 5.2 below shows, Axiom 7 no longer holds.

Example 5.2. Let $S = \{s_1, s_2\}$. Let the constant act $\tilde{1}$ have constant utility 1, so that the minimum weighted expected utility of $\tilde{1}$ is 1 as long as ϕ is normal. Let $p_c \in \Delta(S)$ be the measure such that $p_c(s_1) = c$ for $c \in [0, 1]$. Let ϕ be a confidence function on $\Delta(S)$ such that the confidence value for $p_c \in \Delta(S)$ is

$$\phi(p_c) = \begin{cases} 1, \text{ if } c \geq \frac{1}{2} \\ \frac{1}{2^1}, \text{ if } c \in [\frac{1}{8}, \frac{1}{2}) \\ \frac{1}{2^2}, \text{ if } c \in [\frac{1}{32}, \frac{1}{8}) \\ \dots \\ \frac{1}{2^n}, \text{ if } c \in [\frac{1}{2^{2n+1}}, \frac{1}{2^{2n-1}}), \text{ for } n \in \mathbb{N}. \end{cases}$$

Clearly, ϕ is normal, since $\phi(p_{\frac{1}{2}}) = 1$. It is also easy to see from the definition that ϕ is weakly* upper semicontinuous. Lastly, to check quasi-concavity, note that a function which is nondecreasing up to a point and is nonincreasing from that point on is quasiconcave. Therefore ϕ is quasi-concave.

We describe the utility of an act f on a state space $S = \{s_1, \dots, s_n\}$ using a *utility profile* with the format $(u(f(s_1)), \dots, u(f(s_n)))$. Consider the sequence of acts $\{f_n\}_{n\geq 1}$ with utility profiles as follows

$$f_1 = \left(2, \frac{2}{7}\right)$$
$$f_2 = \left(4, \frac{4}{31}\right)$$
$$f_3 = \left(8, \frac{8}{127}\right)$$
$$\dots$$
$$f_n = \left(2^n, \frac{2^n}{2^{2n+1}-1}\right).$$

Suppose, by way of contradiction, that there is a fixed $\delta \in \mathbb{R}$ such that $\succeq_\phi^+$ satisfies Axiom 7. In Appendix B, we show that for all $n \geq 1$, $f_n \sim_\phi^+ \tilde{1}$.

Now let $\tilde{m}$ be a constant act with constant utility m. The act $\frac{1}{2}f_n + \frac{1}{2\delta}\tilde{\delta}$ has utility $2^{n-1} + \frac{1}{2}$ in state s_1 and utility $\frac{2^{n-1}}{2^{2n+1}-1} + \frac{1}{2}$ in state s_2. If $c \in [\frac{1}{2^{2m+1}}, \frac{1}{2^{2m-1}})$ for $m \geq 1$, then the weighted expected utility of $\frac{1}{2}f_n + \frac{1}{2\delta}\tilde{\delta}$ with respect to p_c is at least $2^{n-m-2} + 2^{m-2}$. This means that if $n \geq 4 + 2\log_2 \delta$, then the minimum weighted expected utility of $\frac{1}{2}f_n + \frac{1}{2\delta}\tilde{\delta}$ is strictly greater than δ. The details are worked out in Appendix B.

On the other hand, the minimum weighted expected utility of $\frac{1}{2}\tilde{1} + \frac{1}{2}\tilde{\delta}$ is $\frac{1}{2}(1+\delta) < \delta$ for $\delta \geq 1$. Thus, $\frac{1}{2}f_n + \frac{1}{2}\frac{1}{\delta}\tilde{\delta} \succ_{t,\phi}^+ \frac{1}{2}\tilde{1} + \frac{1}{2}\tilde{\delta}$ for sufficiently large n, violating Axiom 7 with $x_* = \tilde{0}$. Although Axiom 7 is violated, it is easy to see that Axioms 1–6 hold. Indeed, as we show, we can get a representation theorem for Axioms 1–6.

5.2 Maxmin weighted expected utility

It is useful to think of the CF model not as probability measures accompanied by confidence values, but rather as a set of "super-probability measures." By *super-probability measure* we mean that by multiplying a probability measure by a positive scalar in $[1, \infty)$, we get a scaled positive vector whose components may sum up to more than 1. A super-probability measure is therefore a nonnegative vector whose components sum to at least 1. This notion is analogous to the *sub-probability measures* used in our previous work on minimax weighted expected regret [7], where a sub-probability measure is a nonnegative vector whose components sum to at most 1. Intuitively, a sub-probability measure is obtained by multiplying a probability measure by a scalar weight that is at most 1. We are also interested in sets containing both super and sub-probability measures. We will call these sets of *generalized probability measures.*

It is often helpful to consider the set of generalized probability measures *supporting* the weighting function. For generalized probability measures p and p', let $p' \geq p$ if for all $s \in S, p'(s) \geq p(s)$.

Definition 5.3 (Upper Support)**.** *The* upper support *of a nonnegative weighting function* $t \circ \phi$ *is the set* $\overline{V}_{t\circ\phi} = \{p' : \exists p(\phi(p) > 0$ *and* $p' \geq t(\phi(p)))\}$.

The upper support of $t \circ \phi$ contains the set of generalized probabilities $t(\phi(p))p$, as well as all generalized probabilities that are larger. Including these larger generalized probabilities does not change the underlying preferences of the upper support, since these larger generalized probabilities will never provide minimum expected utilities. While adding larger generalized probabilities does not affect the minimum expected utility, working with the upper support turns out to be technically convenient, as we shall see.

Define a relation $\succeq_{\overline{V}_{t\circ\phi}}$ by taking

$$f \succeq_{\overline{V}_{t\circ\phi}} g \Leftrightarrow \inf_{p \in \overline{V}_{t\circ\phi}} \int_S u(f)dp \geq \inf_{p \in \overline{V}_{t\circ\phi}} \int_S u(g)dp.$$

Just as before, we can convert back and forth between the upper support of a weighting function and the weighting function itself. The proof is analogous to that for Proposition 4.4 and is left to the reader.

Proposition 5.4. $\succeq_{\overline{V}_{t\circ\phi}} = \succeq^+_{t,\phi}$.

For the results beyond this point, we assume that the state space S is finite, since we make use of results due to Halpern and Leung [7], which are proved under the assumption of a finite state space.

Theorem 5.5. *Let* $t : (0, 1] \to \mathbb{R}^+$ *be a strictly decreasing function with* $t(1) > 0$*. For all* X*, nonconstant* U^+*,* S*, and normal* ϕ*, the preference order* $\succeq^+_{t,\phi}$ *satisfies Axioms 1–6. Furthermore, if* t *is continuous,* $\lim_{x\to 0^+} t(x) = \infty$*, and the preference order* $\succeq$ *on the acts in* $\mathcal{F}$ *satisfies Axioms 1–6, then there exists a a nonnegative utility function* U^+ *on* X *and a regular* fuzzy confidence function* $\phi : \Delta(S) \to [0, 1]$ *such that* $t \circ \phi$ *has convex upper support, and* $\succeq = \succeq^+_{t,\phi}$*. Moreover,* U^+ *is unique up to positive linear transformations, and* ϕ *is unique in the sense that if* ϕ' *is such that* $\succeq^+_{t,\phi'} = \succeq$ *and* $\phi' \circ t$ *has convex upper support, then* $\phi = \phi'$*.*

Theorem 5.5 characterizes t-maxmin weighted expected utility without the threshold α_0 of CF. By doing so, we show that the lower bound α_0 on the confidence or weight of probabilities is not a crucial part of the characterization of a weighted version of MMEU. Moreover, we provide a uniqueness result that is in some sense stronger than that by CF [3], in that our uniqueness result directly identifies a "representative" set of beliefs, while the CF construction [3] needs to be maximal in order to be unique. For example, consider a state space S with two states, and the regular* fuzzy set ϕ such that $\phi(p) = 1$ for all $p \in \Delta(S)$. Consider a second regular* fuzzy set ϕ' where $\phi'(p) = \frac{1}{1+\min_{s\in S} p(s)}$. It is not difficult to check that both sets induce the same maxmin preferences in the Chateaneuf and Faro representation, since the supports of the two regular* fuzzy sets have the same convex hull.

The requirement that $\lim_{x\to 0^+} t(x) = \infty$ is necessary to model probability measures that are arbitrarily close to being "ignored". This requirement was not necessary in the representation that made use of a lower bound α_0. However, there is another natural way to relax the constraints on t without introducing a lower bound α_0. As we show in the next section, if instead of restricting to nonnegative utilities, we restrict to nonpositive utilities, then we can drop the requirement that $\lim_{x\to 0^+} t(x) = \infty$, thus allowing a larger set of transformation functions.

5.3 Nonpositive utilities

Although the preceding results provide a relatively simple characterization of t-weighted maxmin expected utility, we have not yet presented the full picture. In the preceding results, just as in the CF model [3], we have restricted utilities of acts to be nonnegative. It is easy to see why the restriction to nonnegative utilities was necessary. A larger weight makes positive utilities better but negative utilities worse. If we were to allow utilities to range over positive and negative values, the resulting decision rule would have very different, rather unintuitive behavior.

It turns out we can get a simpler decision rule, characterized almost exactly[1] by Axioms 1–6, if we look at *nonpositive utilities* instead of nonnegative utilities; in this section, we consider a representation that is restricted to *nonpositive* utilities, rather than nonnegative utilities. We use the notation U^- to indicate a nonpositive utility function.

Definition 5.6 (Weighted maxmin representation)**.** *Given a confidence function $\phi : \Delta(S) \to [0, 1]$ and strictly increasing transformation function $t : [0, 1] \to \mathbb{R}^+$, define $\succeq^-_{t,\phi}$ as follows:*

$$f \succeq^-_{t,\phi} g \Leftrightarrow \min_{p\in\Delta(S)} t(\phi(p)) \sum_{s\in S} p(s)u(f, s) \geq \min_{p\in\Delta(s)} t(\phi(p)) \sum_{s\in S} p(s)u(g, s).$$

The $^-$ superscript on $\succeq^-_{t,\phi}$ denotes that the relation is defined on acts with nonpositive utilities. One benefit of using nonpositive utilities

[1]Because we restrict to nonpositive utilities instead of nonnegative utilities, instead of a worst outcome/act we now have a best outcome/act instead. Thus Axiom 4 no longer holds and is replaced by Axiom 8.

instead of nonnegative utilities is that we no longer need to transform confidence values $\phi(p)$ in $(0,1]$ into multiplicative weights $t(\phi(p)) \in [0,\infty)$. Instead, because a larger multiplicative confidence value results in utilities that are more negative, we can simply use the confidence function as the weights. Equivalently, we can take t to be the identity. Arguably this is the most natural choice for t, and minimizes concerns regarding which transformation function to use.

We show that preferences generated by the weighted maxmin representation is characterized by Axioms 1–6, with Axiom 4 replaced by the following axiom:

Axiom 8 (Best Act Independence)**.** *There exists a best outcome $\overline{x} \in X$ such that $\overline{x}^* \succeq f$ for every $f \in \mathcal{F}$. Moreover,*

$$f \sim g \Rightarrow \alpha f + (1-\alpha)\overline{x}^* \sim \alpha g + (1-\alpha)\overline{x}^*.$$

In the case of nonpositive utilities, as is in the case of minimax weighted expected regret (MWER) [7], it is useful to look at the *lower support* $\underline{V}_{t\circ\phi}$ formed by the set of sub-probabilities, defined by

$$\underline{V}_{t\circ\phi} = \{p' : \exists p(p' \leq t(\phi(p))p)\}.$$

Theorem 5.7. *Let $t : [0,1] \to \mathbb{R}^+$ be a strictly increasing, continuous transformation such that $t(1) > 0 \geq t(0)$. For all X, nonconstant U^-, S, and regular* fuzzy ϕ, the preference order $\succeq^-_{t,\phi}$ satisfies Axioms 1–3, 5–6, and 8. Conversely, if a preference order $\succeq$ on the acts in $\mathcal{F}$ satisfies Axioms 1–3, 5–6, and 8, then there exists a nonpositive utility function U^- on X and a confidence function $\phi : \Delta(S) \to [0,1]$ such that ϕ is regular* fuzzy, has convex lower support, and $\succeq = \succeq^-_{t,\phi}$. Moreover, U^- is unique up to positive linear transformations, and ϕ is unique in the sense that if ϕ' is such that $\succeq^-_{t,\phi'} = \succeq$ and $\phi \circ t$ has convex lower support, then $\phi = \phi'$.*

Note that the transformation t in Theorem 5.7 has domain $[0,1]$ instead of $(0,1)$. This is because in a setting with nonpositive utilities, a confidence value of 0 can be mapped to a weight of 0, contributing nothing to the definition of the preferences. This is analogous to a measure being ignored in the case of nonnegative utilities. Furthermore, t is required to be strictly increasing, instead of decreasing, since a larger multiplier amplifies the significance of a negative utility value. We need that $t(1) > 0$, since if $t(1) = 0$ then the preferences will be trivial. In the second part of the theorem, we need $t(0) \leq 0$ in order to find a representation for all possible preferences that satisfy the axioms. For example, suppose the preference $\succeq$ is such that $(c,0) \sim (c',0)$ for all $c, c' \in \mathbb{R}-$. Intuitively, this means that the first state is ignored. More precisely, any probability measure giving positive probability to the first state should be ignored. If $t(0) > 0$, then we do not have the representation power to ignore these probability measures. Therefore, we are unable to find a representation for $\succeq$.

5.4 The case of general acts

We have considered two different settings, one restricted to nonnegative utilities, and one restricted to nonpositive utilities. One might wonder whether a maxmin weighted expected utility representation could apply to a setting that include both positive and negative utilities. Recall that in the case of nonnegative utilities, a large positive multiplier on the utility *decreases* the impact of the constraint or weighted probability measure, while in the case of nonpositive utilities, a large positive multiplier on the utility *increases* the impact of the constraint or weighted probability measure. As a result, to have reasonable behavior when dealing with both positive and negative utilities, the multiplier on a utility value must depend not only on the probability measure, but also on the utility value itself (whether it is positive or negative).

6 Acknowledgments

The authors thank Leandro Chaves Rêgo and the Theory and Decision reviewers for their insights and useful comments.

A Proof of Theorem 4.2

Proof of Theorem 4.2 . We assume that t is continuous and strictly decreasing, and that there exists some $\beta > 0$ such that $[\beta, \beta/\alpha_0^*(\phi)] \in range(t)$. Recall that $\alpha_0^*(\phi) = \max\{\alpha_0, \min_{p\in\Delta(S)} \phi(p)\}$.

Let $\alpha_0' = t^{-1}(\frac{\beta}{\alpha_0^*})$ and for all $p \in \Delta(S)$, let $\phi'(p) = t^{-1}(\frac{\beta}{\phi(p)})$. It is easy to see that, for all acts f, g,

$$\begin{aligned}
&\min_{p\in L_{\alpha_0}\phi} \frac{1}{\phi(p)} \int_S u(f)dp \geq \min_{p\in L_{\alpha_0}\phi} \frac{1}{\phi(p)} \int_S u(g)dp \\
\text{iff } &\min_{\{p:\phi(p)\geq\alpha_0'\}} \frac{\beta}{\phi(p)} \int_S u(f)dp \geq \min_{\{p:\phi(p)\geq\alpha_0'\}} \frac{\beta}{\phi(p)} \int_S u(g)dp \\
\text{iff } &\min_{p\in L_{\alpha_0'}\phi'} t(\phi'(p)) \int_S u(f)dp \geq \min_{p\in L_{\alpha_0'}\phi'} t(\phi'(p)) \int_S u(g)dp,
\end{aligned}$$

since for all $p \in L_{\alpha_0}\phi$,

$$\frac{\beta}{\phi(p)} = t(t^{-1}(\frac{\beta}{\phi(p)})) = t(\phi'(p)).$$

Now we show that if $t(1) = \beta$, then ϕ' must be a regular* fuzzy set. Since ϕ is normal, there exists p^* such that $\phi(p^*) = 1$. By definition of ϕ', $\phi'(p^*) = t^{-1}(\frac{\beta}{\phi(p^*)}) = t^{-1}(\beta) = 1$, so ϕ' is normal.

To show that ϕ' is weakly* upper semicontinuous, we must show that the set $L_\alpha\phi' = \{p \in \Delta(S) : \phi'(p) \geq \alpha\}$ is weakly* closed for $\alpha \in [0, 1]$. In other words, we have to show that the set $L_\alpha\phi'$ contains all of its limit points, for all $\alpha \in [0, 1]$. Now, for $\alpha = 0$, $L_\alpha\phi' = \Delta(S)$ and is closed. So consider the case $\alpha > 0$.

Recall from our definition of ϕ' that $\phi'(p) = t^{-1}(\frac{\beta}{\phi(p)})$ for all p. Suppose $p_n \to p$. Observe that $\frac{\beta}{\phi(p_n)}$ is in the domain of t^{-1} for all n, since $[\beta, \beta/\alpha_0^*(\phi)] \in range(t)$. Note that for all p, $\phi'(p) \geq \alpha$ if and only if

$$\begin{aligned}
&t^{-1}(\frac{\beta}{\phi(p)}) \geq \alpha \\
\text{iff } &\frac{\beta}{\phi(p)} \leq t(\alpha) \\
\text{iff } &\phi(p) \geq \frac{\beta}{t(\alpha)},
\end{aligned}$$

where $t(\alpha) \geq \beta$ since $0 < \alpha \leq 1$, t is monotonically decreasing, and $t(1) = \beta$. Since ϕ is assumed to be weakly* upper semicontinuous, and $\phi(p_n) \geq \frac{\beta}{t(\alpha)}$ for all n, we have $\phi(p) \geq \frac{\beta}{t(\alpha)}$. Therefore, $\phi'(p) \geq \alpha$, as required.

Finally, to show that ϕ' is quasi-concave, let $\gamma \in [0, 1]$. Using the fact that t is strictly monotonically decreasing, we have that

$$\begin{aligned}
&\phi(\gamma p_1 + (1-\gamma)p_2) \geq \min(\phi(p_1), \phi(p_2)) \\
\Rightarrow\ &\frac{\beta}{\phi(\gamma p_1 + (1-\gamma)p_2)} \leq \max(\frac{\beta}{\phi(p_1)}, \frac{\beta}{\phi(p_2)}) \\
\Rightarrow\ &t^{-1}(\frac{\beta}{\phi(\gamma p_1 + (1-\gamma)p_2)}) \geq \min(t^{-1}(\frac{\beta}{\phi(p_1)}), t^{-}(\frac{\beta}{\phi(p_2)})) \\
\Rightarrow\ &\phi'(\gamma p_1 + (1-\gamma)p_2) \geq \min(\phi'(p_1), \phi'(p_2)).
\end{aligned}$$

For the other direction, suppose that $t(1) = \beta$ and that ϕ' is a regular* fuzzy confidence function. We want to show that ϕ defined by $\phi(p^*) = \frac{1}{t(\phi'(p^*))}$ is also regular* fuzzy. The arguments for this direction are analogous to those used to show the first direction.

□

B Details of Example 5.2

We now show that for all $n \geq 1$, $f_n \sim_\phi^+ \tilde{1}$.

Suppose $c \in [\frac{1}{2^{2m+1}}, \frac{1}{2^{2m-1}})$. The weighted expected utility of f_n with respect to p_c is

$$2^m \left[c2^n + (1-c)\frac{2^n}{2^{2n+1}-1} \right], \text{ for } m \in \{0, 1, 2, \ldots\}.$$

If $m = n$, note that

$$2^n \left[\frac{1}{2^{2n+1}}2^n + \frac{2^{2n+1}-1}{2^{2n+1}}\frac{2^n}{2^{2n+1}-1} \right] = 1.$$

If $m < n$, then

$$\begin{aligned} 2^m \left[c2^n + (1-c)\frac{2^n}{2^{2n+1}-1} \right] &\geq 2^m \left[\frac{1}{2^{2m+1}}2^n + \frac{2^{2m-1}-1}{2^{2m-1}}\frac{2^n}{2^{2n+1}-1} \right] \\ &= \frac{2^n}{2^{m+1}} + \frac{2^{2m-1}-1}{2^{m-1}}\frac{2^n}{2^{2n+1}-1} \\ &\geq \frac{2^n}{2^{m+1}} \geq 1. \end{aligned}$$

If $m > n$, then

$$\begin{aligned} 2^m \left[c2^n + (1-c)\frac{2^n}{2^{2n+1}-1} \right] &\geq 2^m \left[\frac{1}{2^{2m+1}}2^n + \frac{2^{2m-1}-1}{2^{2m-1}}\frac{2^n}{2^{2n+1}-1} \right] \\ &= \frac{2^n}{2^{m+1}} + \frac{2^{2m-1}-1}{2^{m-1}}\frac{2^n}{2^{2n+1}-1} \\ &\geq \frac{2^n}{2^{m+1}} + \frac{2^{2m-1}-1}{2^{m-1}}\frac{1}{2^{n+1}} \\ &\geq \frac{2^n}{2^{m+1}} + \frac{2^{2m-1}}{2^{m+n}} - \frac{1}{2^{m+n}} \\ &\geq 2^{m-n-1} \geq 1. \end{aligned}$$

If $c \in [\frac{1}{2}, 1]$, then the weighted expected utility of f_n is

$$c2^n + (1-c)\frac{2^n}{2^{2n+1}-1} \geq c2^n \geq 2^{n-1}.$$

Therefore, for all n, the minimum weighted expected utility of f_n is 1, so $f_n \sim^+_\phi \tilde{1}$.

Now let $\tilde{m}$ be a constant act with constant utility m. The act $\frac{1}{2}f_n + \frac{1}{2\delta}\tilde{\delta}$ has utility $2^{n-1} + \frac{1}{2}$ in state s_1 and utility $\frac{2^{n-1}}{2^{2n+1}-1} + \frac{1}{2}$ in state s_2. If $c \in [\frac{1}{2^{2m+1}}, \frac{1}{2^{2m-1}})$ for $m \geq 1$, then the weighted expected utility of $\frac{1}{2}f_n + \frac{1}{2\delta}\tilde{\delta}$ with respect to p_c is

$$\begin{aligned} &2^m \left[c\left(2^{n-1} + \frac{1}{2}\right) + (1-c)\left(\frac{2^{n-1}}{2^{2n+1}-1} + \frac{1}{2}\right) \right] \\ \geq &\frac{1}{2^{m+1}}\left(2^{n-1}\right) + \frac{2^{2m-1}-1}{2^{m-1}}\left(\frac{1}{2}\right) \\ \geq &2^{n-m-2} + 2^{m-2}. \end{aligned}$$

Suppose that $n \geq 4 + 2\log_2 \delta$ and $\delta \geq 1$. If $n \geq m + 2 + \log_2 \delta$,

$$2^{n-m-2} + 2^{m-2} > 2^{\log_2 \delta} = \delta.$$

Otherwise, if $n < m + 2 + \log_2 \delta$, since $n \geq 4 + 2\log_2 \delta$, it follows that $m \geq \log_2 \delta + 2$, and

$$2^{n-m-2} + 2^{m-2} > 2^{\log_2 \delta} = \delta.$$

If $c \geq \frac{1}{2}$, then the weighted expected utility of $\frac{1}{2} f_n + \frac{1}{2\delta}\tilde{\delta}$ with respect to p_c is

$$\begin{aligned} & c\left(2^{n-1} + \frac{1}{2}\right) + (1-c)\left(\frac{2^{n-1}}{2^{2n+1}-1} + \frac{1}{2}\right) \\ > & \frac{1}{2}2^{n-1} \geq \frac{1}{2}2^{3+2\log_2 \delta} \geq 2^2\delta^2 > \delta, \end{aligned}$$

since $\delta \geq 1$. This means that if $n \geq 4 + 2\log_2 \delta$, then the minimum weighted expected utility of $\frac{1}{2} f_n + \frac{1}{2\delta}\tilde{\delta}$ is strictly greater than δ.

C Proof of Theorem 5.5

We show here that if a family of preferences $\succeq$ satisfies Axioms 1–6, then $\succeq$ can be represented as maximizing weighted expected utility with respect to a regular confidence function and a utility function. We make use of many of the same techniques as used in [7]. Key differences are highlighted.

First, we establish a von-Neumann-Morgenstern expected utility function over constant acts. This part follows the CF proof, rather than the proof in [7].

Lemma C.1. *If Axioms 1, 3 and 5 hold, then there exists a nonconstant function $U : X \to \mathbb{R}$, unique up to positive affine transformations, such that for all constant acts l^* and $(l')^*$,*

$$l^* \succeq (l')^* \Leftrightarrow \sum_{\{y:\, l^*(y)>0\}} l(y)U(y) \geq \sum_{\{y:\, l'(y)>0\}} l'(y)U(y).$$

Proof. As noted by CF, it was shown by Herstein and Milnor [**?**] that Axioms 1, 3 and 5 are sufficient to satisfy the premises of the von-Neumann-Morgenstern theorem. □

Since U is nonconstant, we can choose a U such that the minimum value that it takes on is 0 (for some constant act), and the maximum value it takes on is at least 1. If c is the utility of some lottery l_c, let l_c^* be a constant act such that $l^*(s) = l_c$, so that $u(l_c^*) = c$. The following lemma, whose proof is given in [7] (Lemma 2), follows from Lemma C.1.

Lemma C.2. *$u(l^*_c) \geq u(l^*_{c'})$ iff $l^*_c \succeq l^*_{c'}$; similarly, $u(l^*_c) = u(l^*_{c'})$ iff $l^*_c \sim l^*_{c'}$, and $u(l^*_c) > u(l^*_{c'})$ iff $l^*_c \succ l^*_{c'}$.*

In [7] a slightly different continuity axiom (Axiom 9) is used.

Axiom 9 (Mixture Continuity)**.** *If $f \succ g \succ h$, then there exist $q, r \in (0,1)$ such that*

$$qf + (1-q)h \succ g \succ rf + (1-r)h.$$

It is not difficult to derive Mixture Continuity from completeness (Axiom 1) and Axiom 3. Therefore, from here on, we assume that the preference order satisfies Mixture Continuity.

We establish some useful notation for acts and utility acts (real-valued functions on S). Given a utility act b, let f_b, the act corresponding to b, be the act such that $f_b(s) = l_{b(s)}$, if such an act exists. Conversely, let b_f, the utility act corresponding to the act f, be defined by taking $b_f(s) = u(f(s))$. Note that monotonicity implies that if $f_b = g_b$, then $f \sim g$. That is, only utility acts matter. If c is a real, we take c^* to be the constant utility act such that $c^*(s) = c$ for all $s \in S$.

C.1 Defining a functional on utility acts

Our proof uses the same technique as that used in [7]. Specifically, like Gilboa and Schmeidler [6], we define a functional I on utility acts such that the preference order on utility acts is determined by their value according to I (see Lemma C.4). Using I, we can then determine the weight of each probability in $\Delta(S)$, and prove the desired representation theorem.

Recall that u represents $\succeq$ on constant acts, and that only utility acts matter to $\succeq$. The space of all nonnegative utility acts is the set $\mathcal{B}^+$ of real-valued functions b on S where $b(s) \geq 0$ for all $s \in S$. We now define a functional I on utility acts in $\mathcal{B}^+$ such that for all f, g with $b_f, b_g \in \mathcal{B}^+$, we have $I(b_f) \geq I(b_g)$ iff $f \succeq g$. Let

$$R_f = \{\alpha' : l^*_{\alpha'} \preceq f\}.$$

If $0^* \leq b \leq 1^*$, then f_b exists, and we define

$$I(b) = \sup(R_{f_b}).$$

For the remaining utility acts $b \in \mathcal{B}^+$, we extend I by homogeneity. Let $||b|| = |\max_{s \in S} b(s)|$. Note that if $b \in \mathcal{B}^+$, then $0^* \leq b/||b|| \leq 1^*$, so we define

$$I(b) = ||b|| I(b/||b||).$$

It is worth noting that while in [7] I was extended from the nonpostive utility acts to the entire set of real-valued acts in order to invoke a separating theorem for Banach spaces, the extension is not performed here. Consequently, we will be using a different separating hyperplane theorem than in [7].

Lemma C.3. *If $b_f \in \mathcal{B}^+$, then $f \sim l^*_{I(b_f)}$.*

Proof. Suppose that $b_f \in \mathcal{B}^+$ and, by way of contradiction, that $l^*_{I(b_f)} \prec f$. If $f \sim l^*_0$, then it must be the case that $I(b_f) = 0$, since $I(b_f) \geq 0$ by definition of sup, and $f \sim l^*_0 \prec l^*_\epsilon$ for all $\epsilon > 0$ by Lemma C.2, so $I(b_f) < \epsilon$ for all $\epsilon < 0$. Therefore, $f \sim l^*_{I(b_f)}$. Otherwise, since $b_f \in \mathcal{B}^+$, by monotonicity, we must have $l^*_0 \prec f$, and thus $l^*_0 \prec f \prec l^*_{I(b_f)}$. By mixture continuity, there is some $q \in (0,1)$ such that $q \cdot l^*_0 + (1-q) \cdot l^*_{I(b_f)} \sim l_{(1-q)I(b_f)} \succ f$, contradicting the fact that $I(b)$ is the least upper bound of R_f.

If, on the other hand, $l^*_{I(b_f)} \succ f$, then $l^*_{I(b_f)} \succ f \succeq l^*_{\underline{c}}$, where the existence of $l^*_{\underline{c}}$ is guaranteed by Axiom 4. If $f \sim l^*_{\underline{c}}$ then it must be the case that $I(b_f) = \underline{c}$. This is because $I(b_f) \geq \underline{c}$ since $l^*_{\underline{c}} \succeq l^*_{\underline{c}}$, and $I(b_f) \leq \underline{c}$ since for all $c' > \underline{c}$, $l^*_{c'} \succ f \sim l^*_{\underline{c}}$.

Otherwise, $l^*_{I(b_f)} \succ f \succ l^*_{\underline{c}}$, and by Axiom 3, there is some $q \in (0,1)$ such that $q \cdot l^*_{I(b_f)} + (1-q) l^*_{\underline{c}} \prec f$. Since $qI(b_f) + (1-q)\underline{c} > I(b_f)$, this contradicts the fact that $I(b_f)$ is an upper bound of R_f. Therefore, it must be the case that $l^*_{I(b_f)} \sim f$. □

We can now show that I has the required property.

Lemma C.4. *For all acts f, g such that $b_f, b_g \in \mathcal{B}^+$, $f \succeq g$ iff $I(b_f) \geq I(b_g)$.*

Proof. Suppose that $b_f, b_g \in \mathcal{B}^+$. By Lemma C.3, $l^*_{I(b_f)} \sim f$ and $g \sim l^*_{I(b_g)}$. Thus, $f \succeq g$ iff $l^*_{I(b_f)} \succeq l^*_{I(b_g)}$, and by Lemma C.2, $l^*_{I(b_f)} \succeq l^*_{I(b_g)}$ iff $I(b_f) \geq I(b_g)$. □

We show that the axioms guarantee that I has a number of standard properties. The proof of each property is analogous to its counterpart in [7], but here we deal with nonnegative utility acts, as opposed to nonpositive utility acts.

Lemma C.5. *(a) If $c \geq 0$, then $I(c^*) = c$.*

(b) I satisfies positive homogeneity: if $b \in \mathcal{B}^+$ and $c > 0$, then $I(cb) = cI(b)$.

(c) I is monotonic: if $b, b' \in \mathcal{B}^+$ and $b \geq b'$, then $I(b) \geq I(b')$.

(d) I is continuous: if $b, b_1, b_2, \ldots \in \mathcal{B}^+$, and $b_n \to b$, then $I(b_n) \to I(b)$.

(e) I is superadditive: if $b, b' \in \mathcal{B}^+$, then $I(b + b') \geq I(b) + I(b')$.

Proof. For part (a), if c is in the range of u, then it is immediate from the definition of I and Lemma C.2 that $I(c^*) = c$. If c is not in the range of u, then since $[0,1]$ is a subset of the range of u, we must have $c > 1$, and by definition of I, we have $I(c^*) = |c|I(c^*/|c|) = c$.

For part (b), first suppose that $||b|| \leq 1$ and $b \in \mathcal{B}^+$ (i.e., $0^* \leq b \leq 1^*$). Then there exists an act f such that $b_f = b$. By Lemma C.3, $f \sim l^*_{I(b)}$. We now consider the case that $c \leq 1$ and $c > 1$ separately. If $c \leq 1$, by Worst Independence, $cf_b + (1-c)l^*_0 \sim cl^*_{I(b)} + (1-c)l^*_0$. By Lemma C.4, $I(b_{cf_b+(1-c)l^*_0}) = I(b_{cl^*_{I(b)}+(1-c)l^*_0})$. It is easy to check that $b_{cf_b+(1-c)l^*_0} = cb$, and $b_{cl^*_{I(b)}} + (1-c)l^*_0 = cI(b)^*$. Thus, $I(cb) = I(cI(b)^*)$. By part (a), $I(cI(b)^*) = cI(b)$. Thus, $I(cb) = cI(b)$, as desired.

If $c > 1$, there are two subcases. If $||cb|| \leq 1$, since $1/c < 1$, by what we have just shown $I(b) = I(\frac{1}{c}(cb)) = \frac{1}{c}I(cb)$. Crossmultiplying, we have that $I(cb) = cI(b)$, as desired. And if $||cb|| > 1$, by definition, $I(cb) = ||cb||I(bc/||cb||) = c||b||I(b/||b||)$ (since $bc/||cb|| = b/||b||$). Since $||b|| \leq 1$, by the earlier argument, $I(b) = I(||b||(b/||b||) = ||b||I(b/||b||)$, so $I(b/||b||) = \frac{1}{||b||}I(b)$. Again, it follows that $I(cb) = cI(b)$.

Now suppose that $||b|| > 1$. Then $I(b) = ||b||I(b/||b||)$. Again, we have two subcases. If $||cb|| > 1$, then

$$I(cb) = ||cb||I(cb/||cb||) = c||b||I(b/||b||) = cI(b).$$

And if $||cb|| \leq 1$, by what we have shown for the case $||b|| \leq 1$,

$$I(b) = I(\frac{1}{c}(cb)) = \frac{1}{c}I(cb),$$

so again $I(cb) = cI(b)$.

For part (c), first note that for $b, b' \in \mathcal{B}^+$, if $||b|| \leq 1$ and $||b'|| \leq 1$, then the acts f_b and $f_{b'}$ exist. Moreover, since $b \geq b'$, we must have $(f_b(s))^* \succeq (f_{b'}(s))^*$ for all states $s \in S$. Thus, by Monotonicity, $f_b \succeq f_{b'}$. If either $||b|| > 1$ or $||b'|| > 1$, let $n = \max(||b||, ||b'||)$. Then $||b/n|| \leq 1$ and $||b'/n|| \leq 1$. Thus, $I(b/n) \geq I(b'/n)$, by what we have just shown. By part (b), $I(b) \geq I(b')$.

For part (d), note that if $b_n \to b$, then for all k, there exists n_k such that $b_n - (1/k)^* \leq b_n \leq b_n + (1/k)^*$ for all $n \geq n_k$. Moreover, by the monotonicity of I (part (c)), we have that $I(b - (1/k)^*) \leq I(b_n) \leq I(b + (1/k)^*)$. Thus, it suffices to show that $I(b - (1/k)^*) \to I(b)$ and that $I(b + (1/k)^*) \to I(b)$.

To show that $I(b - (1/k)^*) \to I(b)$, we must show that for all $\epsilon > 0$, there exists k such that $I(b - (1/k)^*) \geq I(b) - \epsilon$. By positive homogeneity (part (b)), we can assume without loss of generality that $||b-(1/2)^*|| \leq 1$ and that $||b|| \leq 1$. Fix $\epsilon > 0$. If $I(b-(1/2)^*) \geq I(b)-\epsilon$, then we are done. If not, then $I(b) > I(b) - \epsilon > I(b - (1/2)^*)$. Since

$||b|| \leq 1$ and $||b - (1/2)^*|| \leq 1$, f_b and $f_{b-(1/2)^*}$ exist. Moreover, by Lemma C.4, $f_b \succ f_{(I(b)-\epsilon)^*} \succ f_{b-(1/2)^*}$. By mixture continuity, for some $p \in (0,1)$, we have $pf_b + (1-p)f_{(b-(1/2)^*} \succ f_{(I(b)-\epsilon)^*}$. It is easy to check that $b_{pf_b+(1-p)f_{b-(1/2)^*}} = b - ((1-p)/2)^*$. Thus, by Lemma C.4, $f_{b-((1-p)/2)^*} \succeq f_{(I(b)-\epsilon)^*}$, and $I(b-((1-p)/2)^*) > I(b)-\epsilon$. Choose k such that $1/k < (1-p)/2$. Then, by monotonicity (part (c)), $I(b - (1/k)^*) \geq I(b - ((1-p)/2)^*) > I(b) - \epsilon$, as desired.

The argument that $I(b + (1/k)^*) \to I(b)$ is similar and left to the reader.

For part (e), if $||b||, ||b'|| \leq 1$, and $I(b), I(b') \neq 0$, consider $\frac{b}{I(b)}$ and $\frac{b'}{I(b')}$. Since $I(\frac{b}{I(b)}) = I(\frac{b'}{I(b')}) = 1$, it follows from Lemma C.3 that $f_{\frac{b}{I(b)}} \sim f_{\frac{b'}{I(b')}}$. By Ambiguity Aversion, for all $p \in (0,1]$, $pf_{\frac{b}{I(b)}} + (1-p)f_{\frac{b'}{I(b')}} \succeq f_{\frac{b}{I(b)}}$. Thus, taking $p = \frac{I(b)}{I(b)+I(b')}$, $I(\frac{b+b'}{I(b)+I(b')}) = \frac{1}{I(b)+I(b')}I(b+b') = I(\frac{I(b)}{I(b)+I(b')}\frac{b}{I(b)} + \frac{I(b')}{I(b)+I(b')}\frac{b'}{I(b')}) \geq I(\frac{b}{I(b)}) = I(\frac{b'}{I(b')}) = 1$. Hence, $I(b + b') \geq I(b) + I(b')$.

If either $||b|| > 1$ or $||b'|| > 1$, and both $I(b) \neq 0$ and $I(b') \neq 0$, then the result easily follows by positive homogeneity (property (b)).

If either $I(b) = 0$ or $I(b') = 0$, let $b_n = b + \frac{1}{n}^*$ and $b'_n = b' + \frac{1}{n}^*$. Clearly $||b_n|| > 0$, $||b'_n|| > 0$, $b_n \to b$, and $b'_n \to b'_n$. By our argument above, $I(b_n + b'_n) \geq I(b_n) + I(b'_n)$ for all $n \geq 1$. The result now follows from continuity.

□

C.2 Defining the confidence function

In this section, we use I to define a confidence function ϕ that maps each $p \in \Delta(S)$ to a confidence value in $[0,1]$. The heart of the proof involves showing that the resulting function ϕ so determined gives us the desired representation.

Given a confidence function ϕ, for $b \in \mathcal{B}^+$, define

$$WE(b) = \inf_{p \in \mathcal{P}} \phi(p)(\sum_{s \in S} b(s)p(s)).$$

Define

$$\underline{E}(b) = \inf_{p \in \mathcal{P}} \sum_{s \in S} b(s)p(s).$$

and

$$E_p(b) = \sum_{s \in S} b(s)p(s).$$

For each probability $p \in \Delta(S)$, define

$$\phi_t(p) = \inf\{\alpha \in \mathbb{R} : I(b) \leq \alpha E_p(b) \text{ for all } b \in \mathcal{B}^+\}, \quad (1)$$

and let $\phi_t(p) = \infty$ if the inf does not exist. Note that $\phi_t(p) \geq 1$, since $E_p((c)^*) = I((c)^*) = c$ for all distributions p and $c \in \mathbb{R}$. Moreover, it is immediate from the definition of $\phi_t(p)$ that $\phi_t(p)E_p(b) \geq I(b)$ for all $b \in \mathcal{B}^+$. The next lemma shows that there exists a probability p where we have equality.

Lemma C.6. *(a) For some distribution p, we have $\phi_t(p) = 1$.*

(b) For all $b \in \mathcal{B}^+$, there exists p such that $\phi_t(p)E_p(b) = I(b)$.

Proof. The proofs of both part (a) and (b) use a separating hyperplane theorem. If U is a convex subset of $\mathcal{B}^+$, and $b \notin U$, then there is a linear functional λ that separates U from b, that is, $\lambda(b') < \lambda(b)$ for all $b' \in U$. We proceed as follows.

For part (a), we must show that there exists a probability measure p such that for all $b \in \mathcal{B}^+$, we have $E_p(b) \geq I(b)$. This would show that $\phi_t(p) = 1$.

Let $U = \{b' \in \mathcal{B}^+ : I(b') \geq 1\}$. U is closed (by continuity of I) and convex (by positive homogeneity and superadditivity of I), and $(0)^* \notin U$. Thus, there exists a linear functional λ such that $\lambda(b') > \lambda((0)^*) = 0$ for $b' \in U$. We can assume without lost of generality that $\lambda(1^*) = 1$.

We want to show that λ is a *positive* linear functional, that is, that $\lambda(b) \geq 0$ if $b \geq 0^*$. Clearly this holds for b' such that $I(b') \geq 1$. If $b' \geq 0^*$, $I(b') < 1$, and $I(b') > 0$, note that $cI(b') = I(cb') \geq 1$ for some $c \geq 0$. Therefore, $I(b') \geq \frac{1}{c} \geq 0$. If $b' \geq 0^*$ and $I(b') = 0$, note that for all $c > 0$, $\lambda(b' + c^*) \geq 0$ by the previous case. Thus, $\lambda(b') \geq 0$. It follows that λ is a positive functional.

Define the probability distribution p on S by taking $p(s) = \lambda(1_s)$. To see that p is indeed a probability distribution, note that since $1_s \geq 0$ and λ is positive, we must have $\lambda(1_s) \geq 0$. Moreover, $\sum_{s \in S} p(s) = \lambda(1^*) = 1$. In addition, for all $b' \in \mathcal{B}$, we have

$$\lambda(b') = \sum_{s \in S} \lambda(1_s)b'(s) = \sum_{s \in S} p(s)b'(s) = E_p(b').$$

Next, we claim that, for $b \in \mathcal{B}^+$,

$$\text{for all } c > 0, \text{ if } I(b) > c, \text{ then } \lambda(b) > c. \tag{2}$$

To see why the claim is true, note that if $I(b) \geq c$, then $I(b/c) \geq 1$ by positive homogeneity, so $\lambda(b/c) \geq 1$ and $\lambda(b) \geq c$. Therefore, $\lambda(b) \geq I(b)$, as desired.

The proof of part (b) is similar to that of part (a). We want to show that, given $b \in \mathcal{B}^+$, there exists p such that $\phi_t(p)E_p(b) = I(b)$. First consider the case where $||b|| \leq 1$. If $I(b) = 0$, then there must exist some s such that $b(s) = 0$, for otherwise there exists $c > 0$ such that $b \geq c^*$, so $I(b) \geq c$. If $b(s) = 0$, let p_s be such that $p_s(s) = 1$. Then $E_{p_s}(b) = 0$, so part (b) of the Lemma holds in this case.

If $||b|| \leq 1$ and $I(b) > 0$, let $U = \{b' : I(b') \geq I(b)\}$. Again, U is closed and convex, and $b \notin U$, so there exists a linear functional λ such that $\lambda(b') > \lambda(b)$ for $b' \in U$. Since $1^* \in U$ and we can assume without loss of generality $\lambda(1^*) = 1$, we must have $\lambda(b) < 1$.

The same argument as that used in the proof of (a) shows that λ is a positive functional.

Therefore, λ determines a probability distribution p such that, for all $b' \in \mathcal{B}^+$, we have $\lambda(b') = E_p(b')$. p, of course, will turn out to be the desired distribution. To show this, we need to show that $\phi_t(p) = I(b)/E_p(b)$. By definition, $\phi_t(p) \geq I(b)/E_p(b)$. To show that $\phi_t(p) \leq I(b)/E_p b$, we must show that $\frac{I(b)}{E_p(b)} \geq \frac{I(b')}{E_p b'}$ for all $b' \in \mathcal{B}^+$. Equivalently, we must show that $I(b)\lambda(b')/\lambda(b) \geq I(b')$ for all $b' \in \mathcal{B}^+$.

Essentially the same argument used to prove (2) also shows that

$$\text{for all } c > 0, \text{ if } \frac{I(b')}{I(b)} \geq c, \text{ then } \frac{\lambda(b')}{\lambda(b)} \geq c.$$

In particular, if $\frac{I(b')}{I(b)} \geq c$, then by positive homogeneity, $\frac{I(b')}{c} \geq I(b)$, so $\frac{b'}{c} \in U$, and $\lambda(\frac{b'}{c}) > \lambda(b)$ and hence $\frac{\lambda(b')}{\lambda(b)} \geq c$.

It follows that $\lambda(b')/(\lambda(b)) \geq I(b')/(I(b))$ for all $b' \in \mathcal{B}^+$. Thus, $I(b)\lambda(b')/\lambda(b) \geq I(b')$ for all $b' \in \mathcal{B}^+$, as required.

Finally, if $||b|| > 1$, let $b' = b/||b||$. By the argument above, there exists a probability measure p such that $\phi_t(p)E_p(b/||b||) = I(b/||b||)$. Since $E_p(b/||b||) = E_p(b)/||b||$, and $I(b/||b||) = I(b)/||b||$, we must have that $\phi_t(p)E_p(b) = I(b)$. □

We can now complete the proof of Theorem 5.5. By Lemma C.6 and the definition of $\phi_t(p)$, for all $b \in \mathcal{B}^+$,

$$I(b) = \inf_{p \in \Delta(S)} \phi_t(p)E_p(b). \tag{3}$$

Recall that, by Lemma C.4, for all acts f, g such that $b_f, b_g \in \mathcal{B}^+$, $f \succeq g$ iff $I(b_f) \geq I(b_g)$. Thus, $f \succeq g$ iff

$$\inf_{p \in \Delta(S)} \left(\phi_t(p) \sum_{s \in S} u(f(s))p(s) \right) \geq \inf_{p \in \Delta(S)} \left(\phi_t(p) \sum_{s \in S} u(g(s))p(s) \right).$$

To get the confidence function ϕ from ϕ_t, note that $\lim_{x \to 0^+} t(x) = \infty$ and $t(1) > 0$. We let $\phi(p) = t^{-1}(t(1)\phi_t(p))$, with the special case $\phi(p) = 0$ if $\phi_t(p) = \infty$. (Note that $t(1)\phi_t(p)$ is in the range of t^{-1}, since $\phi_t(p) \geq 1$, t is nonincreasing, and $\lim_{x \to 0^+} t(x) = \infty$.)

C.3 Properties of the confidence function

In this section, we show that the confidence function ϕ that we constructed satisfies the properties claimed in Theorem 5.5.

We first show that $t \circ \phi = \phi_t$ has convex upper support. To that end, we show that if $c_1 \geq \phi_t(p_1)$ and $c_2 \geq \phi_t(p_2)$, then for all $\alpha \in (0,1)$,

$$(\alpha c_1 p_1 + (1-\alpha) c_2 p_2)(S) \geq \phi_t \left(\frac{\alpha c_1 p_1 + (1-\alpha) c_2 p_2}{(\alpha c_1 p_1 + (1-\alpha) c_2 p_2)(S)} \right).$$

By the definition of ϕ_t, it suffices to show that for all $b \in \mathcal{B}^+$,

$$I(b) \leq (\alpha c_1 p_1 + (1-\alpha) c_2 p_2)(S) E_{\frac{\alpha c_1 p_1 + (1-\alpha) c_2 p_2}{(\alpha c_1 p_1 + (1-\alpha) c_2 p_2)(S)}}(b). \tag{4}$$

It is easy to see that the inequality holds. Let $b \in \mathcal{B}^+$. The right-hand side of (4) is equal to

$$\begin{aligned}
\sum_{s \in S} ((\alpha c_1 p_1(s) + (1-\alpha) c_2 p_2(s)) b(s)) &= \alpha c_1 E_{p_1}(b) + (1-\alpha) c_2 E_{p_2}(b) \\
&\geq \alpha \phi_t(p_1) E_{p_1}(b) + (1-\alpha) \phi_t(p_2) E_{p_2}(b) \\
&\geq \alpha I(b) + (1-\alpha) I(b) \text{ (by (3))} \\
&\geq I(b).
\end{aligned}$$

We now show that ϕ is regular*. Since we've shown that, for some p^*, $\phi_t(p^*) = 1$, we have $\phi(p^*) = t^{-1}(t(1)1) = 1$. Therefore ϕ is normal.

Secondly, we show that ϕ is weakly* upper semicontinuous. We show that if $\{p_n\} \to p$ and $\phi(p_n) \geq \alpha$ for all n, then $\phi(p) \geq \alpha$. Suppose for the purpose of contradiction that $\phi(p) < \alpha$. Then $\phi_t(p) = t(\phi(p)) > t(\alpha)$. By continuity of t, $\phi_t(p_n) = t(\phi(p_n)) > t(\alpha)$ for all sufficiently large n, implying that $\phi(p_n) < \alpha$, contradicting the assumption that $\phi(p_n) \geq \alpha$. Therefore $\phi(p) \geq \alpha$, as required.

We now show that ϕ is quasiconcave; that is, $\phi(\beta p_1 + (1-\beta) p_2) \geq \min\{\phi(p_1), \phi(p_2)\}$ for any $\beta \in [0,1]$. Since t is strictly decreasing, so is t^{-1}. Thus, $-t^{-1}$ is strictly increasing. Moreover, if ϕ_t is quasiconvex then $-t^{-1} \circ \phi_t$ is also quasiconvex. Since the negative of a quasiconvex function is quasiconcave, $t^{-1} \circ \phi_t$ is quasiconcave. Therefore, if we show that ϕ_t is quasiconvex, this would show that $\phi = t^{-1} \circ \phi_t$ is quasiconcave.

Recall from (1) that

$$\phi_t(p) = \inf\{\alpha \in \mathbb{R} : I(b) \leq \alpha E_p(b) \text{ for all } b \in \mathcal{B}^+\}.$$

If $\max\{\phi_t(p_1), \phi_t(p_2)\} \leq c$ for $c \in \mathbb{R}$, then for all $b \in \mathcal{B}^+$, we have

$$I(b) \leq c E_{p_1}(b),$$

and

$$I(b) \leq cE_{p_2}(b).$$

Therefore, for all $b \in \mathcal{B}^+$ and all $\beta \in [0,1]$, by the linearity of $E_p(b)$ with respect to the parameter p,

$$I(b) \leq cE_{\beta p_1 + (1-\beta)p_2}(b).$$

This means that $\phi_t(\beta p_1 + (1-\beta)p_2) \leq c$. Thus, $\phi_t(\beta p_1 + (1-\beta)p_2) \leq \max\{\phi_t(p_1), \phi_t(p_2)\}$. Therefore, ϕ_t is quasiconvex.

C.4 Uniqueness of the representation

In this section, we show that our constructed ϕ is the only regular* fuzzy confidence function such that $t \circ \phi$ has convex upper support, and such that $\succeq^+_{t,\phi} = \succeq$. Our uniqueness result is similar in spirit to the uniqueness results of Gilboa and Schmeidler [6], who show that the convex, closed, and non-empty set of probability measures in their representation theorem for MMEU is unique.

The proof of this result, like the proof of uniqueness in Gilboa and Schmeidler [6], uses a separating hyperplane theorem to show the existence of acts on which two different representations must 'disagree'. The proof presented here is essentially the same as that used in [7], with only superficial changes to accommodate our definitions and notation.

Lemma C.7. *For all confidence functions ϕ', if $\succeq^+_{t,\phi'} = \succeq$ and $t \circ \phi'$ has convex upper support, then $\phi = \phi'$.*

Proof. Suppose for contradiction that there exists a regular* fuzzy confidence function $\phi' \neq \phi$ such that $t \circ \phi'$ has convex upper support, and that $\succeq^+_{t,\phi'} = \succeq^+_{t,\phi}$. Consider the two upper supports $\overline{V}_{t\circ\phi}$ and $\overline{V}_{t\circ\phi'}$. $\overline{V}_{t\circ\phi}$ and $\overline{V}_{t\circ\phi'}$ are both closed. To see why, consider a sequence $\{p_n\}_{n\in\mathbb{N}}$ contained in $p_n \in \overline{V}_{t\circ\phi}$ such that $p_n \to p$. We show that $p \in \overline{V}_{t\circ\phi}$, by showing that for some $q \in \Delta(S)$, $\phi(q) > 0$ and $p \geq t(\phi(q))q$.

We first show that $p \geq t(\phi(q))q$ for some $q \in \Delta(S)$. Recall that for all n, there exists $q_n \in \Delta(S)$ such that $p_n \geq t(\phi(q_n))q_n$. Since $q_n \in \Delta(S)$, $q_{k_m} \to q$ for some subsequence $\{q_{k_m}\}$ and $q \in \Delta(S)$.

Therefore, we have

$$
\begin{aligned}
p &= \lim_{n\to\infty} p_n \\
&\geq \limsup_{n\to\infty} t(\phi(q_n))q_n, \text{ since } p_n \geq t(\phi(q_n))q_n \\
&= \lim_{n\to\infty} \sup_{m\geq n} t(\phi(q_{k_m}))q_{k_m} \\
&= \lim_{n\to\infty} \sup_{m\geq n} t(\phi(q_{k_m})) \lim_{m\to\infty} q_{k_m} \\
&= \lim_{n\to\infty} t(\inf_{m\geq n} \phi(q_{k_m})) \lim_{m\to\infty} q_{k_m}, \text{ since } t \text{ is nonincreasing and continuous} \\
&= t(\liminf_{m\to\infty} \phi(q_{k_m})) \lim_{m\to\infty} q_{k_m}, \text{ by continuity of } t \\
&\geq t(\phi(q))q,
\end{aligned}
$$

since $\phi(q) \geq \limsup_{m\to\infty} \phi(q_{k_m}) \geq \liminf_{m\to\infty} \phi(q_{k_m})$ by upper semicontinuity of ϕ, and t is nonincreasing.

It remains to show that $\phi(q) > 0$. To that end, suppose for the purpose of contradiction that $\phi(q) = 0$. Then it must be the case that $\lim_{m\to\infty} \phi(q_{k_m}) = 0$, since if there exists an $\epsilon > 0$ such that $\lim_{m\to\infty} \phi(q_{k_m}) \geq \epsilon$, then by upper semicontinuity of ϕ it must be the case that $\phi(q) \geq \epsilon$. Since $\lim_{x\to 0^+} t(x) = \infty$, we have that $\lim_{m\to\infty} t(\phi(q_{k_m})) = \infty$. However, recall that $p_n \geq t(\phi(q_n))q_n$ for all n. Since $q_n \in \Delta(S)$ and hence does not vanish, p_n cannot be a convergent sequence. Hence it must be the case that $\phi(q) > 0$.

Therefore, $p \in \overline{V}_{t\circ\phi}$, as required, and that $\overline{V}_{t\circ\phi}$ is closed. The same argument shows that $\overline{V}_{t\circ\phi'}$ is closed.

Without loss of generality, let $q \in \overline{V}_{t\circ\phi'} \setminus \overline{V}_{t\circ\phi}$. Since $\overline{V}_{t\circ\phi}$ and $\{q\}$ are closed, convex, and disjoint, and $\{q\}$ is compact, the separating hyperplane theorem [12] says that there exists $\theta \in \mathbb{R}^{|S|}$ and $c \in \mathbb{R}$ such that

$$\theta \cdot p > c \text{ for all } p \in \overline{V}_{t\circ\phi}, \text{ and } \theta \cdot q < c. \tag{5}$$

By scaling c appropriately, we can assume that $|\theta(s)| \leq 1$ for all $s \in S$. Now we argue that it must be the case that $\theta(s) \geq 0$ for all $s \in S$ (so that θ corresponds to the utility profile of some act with nonnegative utilities). Suppose that $\theta(s') < 0$ for some $s' \in S$. By (5), $\theta \cdot p > c$ for all $p \in \overline{V}_{t\circ\phi}$. Let $p^* \in \overline{V}_{t\circ\phi}$ be any measure with $\phi(p^*) = 1$, and let $p^{**} \in \overline{V}_{t\circ\phi}$ be defined by

$$p^{**}(s) = \begin{cases} p^*(s), \text{ if } s \neq s' \\ \frac{|S| \max\{|c|, \max_{s''\in S} |p^*(s'')|\}}{|\theta(s')|}, \text{ if } s = s'. \end{cases}$$

We have defined p^{**} such that $p^{**} \geq p^*$, since for all $s \in S$, $p^{**}(s) \geq p^*(s)$. To see how, note that $p^{**}(s) = p^*(s)$ for $s \neq s'$, and $p^{**}(s) \geq \max_{s''\in S} |p^*(s'')| \geq p^*(s)$ for $s = s'$. Therefore, p^{**} is in $\overline{V}_{t\circ\phi}$.

Our definition of p^{**} also ensures that $\theta \cdot p^{**} = \sum_{s \in S} p^{**}(s)\theta(s) \leq c$, since

$$\begin{aligned}\sum_{s \in S} p^{**}(s)\theta(s) =& p^{**}(s')\theta(s') + \sum_{s \neq s'} p^{**}(s)\theta(s) \\ \leq& p^{**}(s')\theta(s') + \sum_{s \neq s'} |p^{**}(s)|, \text{ since } |\theta(s)| \leq 1 \\ =& -|S| \max\{|c|, \max_{s'' \in S} |p^*(s'')|\} + \sum_{s \neq s'} |p^{**}(s)| \\ \leq& -|c| \leq c.\end{aligned}$$

This contradicts (5), which says that $\theta \cdot p > c$ for all $p \in \overline{V}_{t \circ \phi}$. Thus it must be the case that $\theta(s) \geq 0$ for all $s \in S$.

Consider the θ given by the separating hyperplane theorem, and let f be an act such that $u \circ f = \theta$. $f \sim l_d^*$ for some constant act l_d^*. Since $\overline{V}_{t \circ \phi}$ and $\overline{V}_{t \circ \phi'}$ as sets of generalized probabilities both represent $\succeq$, and $\overline{V}_{t \circ \phi}$ and $\overline{V}_{t \circ \phi'}$ both contain a normal probability measure,

$$\min_{p \in \overline{V}_{t \circ \phi}} p \cdot (u \circ f) = \min_{p \in \overline{V}_{t \circ \phi}} p \cdot (u \circ l_d^*) = d = \min_{p \in \overline{V}_{t \circ \phi'}} p \cdot (u \circ f).$$

However, by (5),

$$\min_{p \in \overline{V}_{t \circ \phi}} p \cdot (u \circ f) > c > \min_{p \in \overline{V}_{t \circ \phi'}} p \cdot (u \circ f),$$

which is a contradiction. □

D Proof of Theorem 5.7

Proof. The proof is almost the same as the proof of Theorem 5.5. We point out the differences, which are mostly straightforward adaptations from $\mathcal{B}^+$ to $\mathcal{B}^-$. Lemma C.1 and Lemma C.2 hold without change. By Axiom 8, we can assume that the maximum value that u takes on is 0, and by Axiom 1 we can assume that the minimum is no greater than -1.

We now define a functional I on utility acts, as before. All occurrences of $\mathcal{B}^+$ in the proof of Theorem 5.5 needs to be replaced by $\mathcal{B}^-$, defined by the real-valued functions b on S where $b(s) \leq 0$ for all $s \in S$.

More specifically, let

$$R_f = \{\alpha' : l_{\alpha'}^* \preceq f\}.$$

If $0^* \geq b \geq (-1)^*$, then f_b exists, and we define

$$I(b) = \sup(R_{f_b}).$$

For the remaining utility acts $b \in \mathcal{B}^+$, we extend I by homogeneity, as before.

The analog of Lemma C.3 for $b_f \in \mathcal{B}^-$ follows from analogous arguments used in the original proof. The case of $l^*_{I(b_f)} \prec f$, however, is a bit simpler than for the positive case.

Lemma D.1. *If $b_f \in \mathcal{B}^-$, then $f \sim l^*_{I(b_f)}$.*

Proof. Suppose, by way of contradiction, that $l^*_{I(b_f)} \prec f$. If $f \sim l^*_0$, then $I(b_f) \geq 0$ by the definition of I. However, we also have $I(b_f) \leq 0$ by Lemma C.4, so $I(b_f) = 0$, and therefore $f \sim l^*_{I(b_f)}$, as required. Otherwise, $f \prec l^*_0$ by monotonicity, so $l^*_{I(b_f)} \prec f \prec l^*_0$, which, when taken together with mixture continuity, contradicts the definition of I. □

The proof of Lemma C.4 still holds. The analog of Lemma C.5 also follows from simular arguments; we discuss some key differences below.

Lemma D.2. *(a) If $c \leq 0$, then $I(c^*) = c$.*

(b) I satisfies positive homogeneity: if $b \in \mathcal{B}^-$ and $c > 0$, then $I(cb) = cI(b)$.

(c) I is monotonic: if $b, b' \in \mathcal{B}^-$ and $b \geq b'$, then $I(b) \geq I(b')$.

(d) I is continuous: if $b, b_1, b_2, \ldots \in \mathcal{B}^-$, and $b_n \to b$, then $I(b_n) \to I(b)$.

(e) I is superadditive: if $b, b' \in \mathcal{B}^-$, then $I(b + b') \geq I(b) + I(b')$.

Proof. For part (b), instead of making use of Axiom 4 (worst independence), we use Axiom 8 (best independence).

For part (e), note that since $I(b)$ is nonpositive for $b \in \mathcal{B}^-$, $I(\frac{b}{I(b)})$ is not defined, unlike in the case of nonnegative utilities. We use the same proof as in [7]: Clearly, $I(\frac{b}{-I(b)}) = -1$. Therefore, $f_{\frac{b}{-I(b)}} \sim f_{\frac{b'}{-I(b')}} \sim l^*_{-1}$. From Axiom 6 (ambiguity aversion), taking $p = \frac{-I(b)}{-I(b)-I(b')}$, we have

$$I\left(\frac{-I(b)}{-I(b)-I(b')}\frac{b}{-I(b)} + \frac{-I(b')}{-I(b)-I(b')}\frac{b'}{-I(b')}\right) \geq I(\frac{b}{-I(b)}) = -1,$$

which implies that $I(b + b') \geq I(b) + I(b')$, as required. □

We now use I to define a confidence function ϕ. WE, $\underline{E}$, and E are defined as before. For each probability $p \in \Delta(S)$, define

$$\phi_t(p) = \sup\{\alpha \in \mathbb{R} : I(b) \leq \alpha E_p(b) \text{ for all } b \in \mathcal{B}^-\}.$$

Note that $\phi_t(p) \leq 1$, since $E_p((c)^*) = I((c)^*) = c$ for all distributions p and $c \in \mathbb{R}$. Moreover, $\phi_t(p) \geq 0$ for all $b \in \mathcal{B}^-$. The next lemma shows that there exists a probability p where we have equality. The proof of the lemma is similar to that of Lemma C.6, and is left to the reader.

Lemma D.3. *(a) For some distribution p, we have $\phi_t(p) = 1$.*

(b) For all $b \in \mathcal{B}^-$, there exists p such that $\phi_t(p)E_p(b) = I(b)$.

By Lemma D.3 and the definition of $\phi_t(p)$, for all $b \in \mathcal{B}^-$,

$$I(b) = \inf_{p\in\Delta(S)} \phi_t(p)E_p(b).$$

We have $f \succeq g$

$$\text{iff } \inf_{p\in\Delta(S)} \left(\phi_t(p)\sum_{s\in S} u(f(s))p(s)\right) \geq \inf_{p\in\Delta(S)} \left(\phi_t(p)\sum_{s\in S} u(g(s))p(s)\right)$$

$$\text{iff } t(1)\inf_{p\in\Delta(S)} \left(\phi_t(p)\sum_{s\in S} u(f(s))p(s)\right) \geq t(1)\inf_{p\in\Delta(S)} \left(\phi_t(p)\sum_{s\in S} u(g(s))p(s)\right).$$

Since t is strictly increasing, $t(1) > t(0)$. Therefore, since $\phi_t(p) \in [0, 1]$ and $t(0) \leq 0$, $t(1)\phi_t(p)$ is in the range of t, and we can define

$$\phi(p) = t^{-1}(t(1)\phi_t(p)).$$

We now have $f \succeq g$

$$\text{iff } \inf_{p\in\Delta(S)} \left(t(\phi(p))\sum_{s\in S} u(f(s))p(s)\right) \geq \inf_{p\in\Delta(S)} \left(t(\phi(p))\sum_{s\in S} u(g(s))p(s)\right).$$

Finally, uniqueness of the representation follows from arguments analogous to those for nonnegative utilities. □

References

[1] M. Allais. Le comportement de l'homme rationel devant le risque: critique de l'école Americaine. *Econometrica*, 21:503–546, 1953.

[2] F. Anscombe and R. Aumann. A definition of subjective probability. *Annals of Mathematical Statistics*, 34:199–205, 1963.

[3] A. Chateauneuf and J. Faro. Ambiguity through confidence functions. *Journal of Mathematical Economics*, 45:535 – 558, 2009.

[4] G. de Cooman. A behavioral model for vague probability assessments. *Fuzzy Sets and Systems*, 154(3):305–358, 2005.

[5] D. Ellsberg. Risk, ambiguity, and the savage axioms. *The Quarterly Journal of Economics*, 75(4):643–669, 1961.

[6] I. Gilboa and D. Schmeidler. Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics*, 18(2):141–153, 1989.

[7] J. Y. Halpern and S. Leung. Weighted sets of probabilities and minimax weighted expected regret: new approaches for representing uncertainty and making decisions. In *Proc. Twenty-Ninth Conference on Uncertainty in Artificial Intelligence (UAI 2012)*, pages 336–345, 2012.

[8] T. Hayashi. Regret aversion and opportunity dependence. *Journal of Economic Theory*, 139(1):242–268, 2008.

[9] P. Klibanoff, M. Marinacci, and S. Mukerji. A smooth model of decision making under ambiguity. *Econometrica*, 73(6):1849–1892, November 2005.

[10] F. Maccheroni, M. Marinacci, and A. Rustichini. Ambiguity aversion, robustness, and the variational representation of preferences. *Econometrica*, 74(6):1447–1498, 2006.

[11] S. Moral. Calculating uncertainty intervals from conditional convex sets of probabilities. pages 199–206, 1992.

[12] R. T. Rockafellar. *Convex Analysis*. Princeton University Press, Princeton, N. J., 1970.

[13] L. J. Savage. *The Foundations of Statistics*. Wiley, New York, 1954.

[14] P. Walley. Statistical inferences based on a second-order possibility distribution. *International Journal of General Systems*, 26(4):337–383, 1997.