A Hierarchy of Ambiguity Aversion[1]

Lorenz Hartmann* and T. Florian Kauffeldt†
This version: November 24, 2021

Abstract

Ambiguity aversion is commonly associated with a preference for mixing among acts. In this paper, we provide an axiomatic hierarchy that characterizes increasingly strong levels of ambiguity aversion. Each level is specified by k, the maximum number of outcomes of the mix that the decision maker is ensured to accept. For the class of invariant biseparable preferences (Ghirardato et al., 2004), we show that each level k of the hierarchy is equivalent to a specific property of the representation functional: Acts with at most k different outcomes are evaluated at the minimum expected utility w.r.t. the set of dominating measures. Furthermore, the extreme levels of the hierarchy correspond to the prominent definitions of ambiguity aversion introduced in Schmeidler (1989) and Ghirardato and Marinacci (2002). We show that in the Choquet expected utility model the hierarchy only has three levels. Furthermore, our level $k = 2$ axiomatizes exact capacities, providing the first such characterization.

Keywords: Ambiguity Aversion, Preference for Mixing, Invariant Biseparable Preferences, Choquet Expected Utility Preferences, Exact Capacities

Acknowledgements. We thank Alain Chateauneuf, Jürgen Eichberger, Itzhak Gilboa, David Kelsey, Georg Nöldeke, Jörg Oechssler, David Schmei-

[1]This paper is based on the working paper Hartmann and Kauffeldt (2018) and Chapter 2 in Hartmann (2018).

*University of Basel, Faculty of Business and Economics, Peter Merian-Weg 6, 4002 Basel, Switzerland.
Email: lorenz.hartmann@unibas.ch

†Faculty of International Business, Heilbronn University, Germany.
Email: mail@florian-kauffeldt.com

dler, Stefan Trautmann, the seminar participants at the University of Heidelberg, and the conference participants at RUD for helpful comments. Lorenz Hartmann acknowledges the financial support by the Swiss National Science Foundation (SNF) through grant 200915.

1. Introduction

Ellsberg (1961) demonstrated that decision makers (DM) often dislike situations in which probabilities are unknown - a phenomenon that is referred to as *ambiguity aversion*. There is ample empirical evidence supporting Ellsberg's observation (see, e.g., Camerer and Weber, 1992). This highlights the necessity to define ambiguity aversion in terms of preference conditions.

Schmeidler (1989) proposed the first definition of ambiguity aversion according to which a DM is ambiguity averse if she has a *preference for mixing* among acts, the idea being that an ambiguity averse DM is better off when she can average out utility distributions.

Schmeidler's approach was the only definition used in the literature until the contributions of Epstein (1999) and Ghirardato and Marinacci (2002). Both take a comparative approach in the spirit of Yaari (1987): A DM is ambiguity averse if she is more ambiguity averse than some ambiguity neutral benchmark. Epstein suggests probabilistic sophistication (Machina and Schmeidler, 1992) as the benchmark, while Ghirardato and Marinacci suggest subjective expected utility (Savage, 1954) as the benchmark.

In this paper we introduce an axiomatic *hierarchy of ambiguity aversion*. Level k of this hierarchy is characterized by a preference for mixing among indifferent acts whenever the mix has at most k different outcomes. The levels are thus characterized by concrete and increasingly strong conditions on preference for mixing among acts.

Our main result (Theorem 1) shows that for invariant biseparable preferences (Ghirardato et al., 2004) each level of the hierarchy is characterized by a specific property of the representation functional regarding the set of measures that dominate the preference relation.[2] The proof is inspired by Schmeidler (1972), a paper on cooperative game theory. An implication of this result is that the definition of Schmeidler (1989) corresponds to the strongest level and the definition of Ghirardato and Marinacci (2002) to the

[2]The set of dominating measures, defined in equation (1) below, is introduced in Ghirardato and Marinacci (2002) and generalizes the concept of the "Core".

weakest level of the hierarchy. We thus provide a unifying framework for these two influential definitions.

Furthermore, we show that our hierarchy has interesting consequences within the Choquet expected utility model (Schmeidler, 1989). The second level characterizes *exact capacities* (Shapley, 1965), providing the first axiomatization of this kind and a solution to an unsolved problem. In addition, we show that the hierarchy only has three levels, implying that within this important model there is less "room" for different levels of ambiguity aversion compared to the more general class of invariant biseparable preferences.

The existing literature has introduced other interesting approaches to model ambiguity aversion, which are, however, not directly related to this paper. For instance, Klibanoff et al. (2005) introduce a definition that is specific to their smooth ambiguity model.

We point out that Chandrasekher et al. (2021) also introduce a hierarchy of ambiguity aversion. Apart from the identical terminology for the axioms, "k-Ambiguity Aversion", the approaches are very different: Chandrasekher et al. consider mixes that result in constant acts (dubbed *complete hedges*); the parameter k refers to the maximum *number of acts* amongst which the DM mixes. In our paper the parameter k refers to the *number of outcomes* of the resulting mix. The overlap of these two approaches is that our weakest level (1-ambiguity-aversion) and their strongest level ($|S|$-ambiguity-aversion) coincide.

The paper is organized as follows. The next section contains preliminaries and the details of the definitions of ambiguity aversion by Schmeidler (1989) as well as Ghirardato and Marinacci (2002). In section 3, we define our hierarchy of ambiguity aversion. Section 4 presents our results. Section 5 concludes with a summary and discussion. All proofs are in the appendix.

2. Preliminaries

2.1. Decision-theoretic set-up

We use the framework of Anscombe and Aumann (1963). Let S be a finite state space with power set Σ (*the set of events*), Δ the set of probability distributions over S, and X the set of finite-support lotteries (*the set of outcomes*) over some set of prizes. The set of all acts $f : S \to X$ is denoted by $\mathcal{F}$. With the usual abuse of notation we denote by $x \in X$ also the constant act with $x(s) = x$ for all $s \in S$. An act $f \in \mathcal{F}$ is called a *k-act* if it maps to at most k different outcomes. The set of k-acts is denoted by $\mathcal{F}_k$. Thus,

$\mathcal{F}_1 = X$ is the set of constant acts and $\mathcal{F}_2$ is the set of binary acts. Mixtures of acts are performed pointwise: For all $f, g \in \mathcal{F}$ and each $\alpha \in [0,1]$, the act $\alpha f + (1-\alpha)g$ yields $\alpha f(s) + (1-\alpha)g(s) \in X$ for all $s \in S$.

Preferences over $\mathcal{F}$ are represented by a binary relation $\succcurlyeq$ whose asymmetric and symmetric components are denoted by $\succ$ and $\sim$, respectively. A functional $V : \mathcal{F} \to \mathbb{R}$ is called a *representation functional* of $\succcurlyeq$ if for all $f, g \in \mathcal{F}$, $f \succcurlyeq g \iff V(f) \geq V(g)$. Throughout the paper we consider *invariant biseparable* (IB) preferences. These are preferences that satisfy the following five axioms (cf. Ghirardato et al., 2004).

Axiom 1 (Weak Order)**.** *For all $f, g, h \in \mathcal{F}$: (1) either $f \succcurlyeq g$ or $g \succcurlyeq f$, (2) if $f \succcurlyeq g$ and $g \succcurlyeq h$, then $f \succcurlyeq h$.*

Axiom 2 (Certainty Independence)**.** *If $f, g \in \mathcal{F}$, $x \in X$, and $\alpha \in (0,1]$, then $f \succcurlyeq g \Leftrightarrow \alpha f + (1-\alpha)x \succcurlyeq \alpha g + (1-\alpha)x$.*

Axiom 3 (Archimedean)**.** *If $f, g, h \in \mathcal{F}$ and $f \succ g \succ h$, then there are $\alpha, \beta \in (0,1)$ such that $\alpha f + (1-\alpha)h \succ g \succ \beta f + (1-\beta)h$.*

Axiom 4 (Monotonicity)**.** *If $f, g \in \mathcal{F}$ and $f(s) \succcurlyeq g(s)$ for all $s \in S$, then $f \succcurlyeq g$.*

Axiom 5 (Non-Degeneracy)**.** *There are $f, g \in \mathcal{F}$ such that $f \succ g$.*

Ghirardato et al. (2004), Amarante (2009) and Chandrasekher et al. (2021) provide representation results for IB preferences. When Axiom 2 is strengthened to the standard independence axiom,[3] an axiomatization of *subjective expected utility* (SEU) preferences is achieved (Anscombe and Aumann, 1963): There exists a unique prior $P \in \Delta$ and an affine utility function $u : X \to \mathbb{R}$ such that $f \succcurlyeq g \iff \int u(f)dP \geq \int u(g)dP$ for all $f, g \in \mathcal{F}$.

Choquet expected utility (CEU) preferences, introduced by Schmeidler (1989), are a special case of IB preferences. Schmeidler introduces the axiom Comonotonic Independence, a strengthening of Certainty Independence:[4]

Axiom 6 (Comonotonic Independence)**.** *For all pairwise comonotonic acts $f, g, h \in \mathcal{F}$ and $\alpha \in (0,1]$, $f \succcurlyeq g \iff \alpha f + (1-\alpha)h \succcurlyeq \alpha g + (1-\alpha)h$.*

[3]If $f, g, h \in \mathcal{F}$ and $\alpha \in (0,1]$, then $f \succcurlyeq g \Leftrightarrow \alpha f + (1-\alpha)h \succcurlyeq \alpha g + (1-\alpha)h$.

[4]Two acts f and g are called *comonotonic* if there are no $s, s' \in S$ such that $f(s) \succ f(s')$ and $g(s) \prec g(s')$.

Schmeidler shows that preferences that satisfy the axioms 1, 3, 4, 5 and 6 can be represented by a unique normalized and monotonic set function (*capacity*) $\nu : \Sigma \to [0,1]$ and a non-constant affine utility function $u : X \to \mathbb{R}$ such that

$$f \succcurlyeq g \text{ if and only if } \int u(f)\, d\nu \geq \int u(g)\, d\nu,$$

where the integral is the Choquet integral (Choquet, 1954) with respect to ν.

Let $V : \mathcal{F} \to \mathbb{R}$ be a representation functional of an IB preference relation $\succcurlyeq$ and define $u(x) \equiv V(x)$ for all $x \in X$. The following *set of dominating measures* is introduced in Ghirardato and Marinacci (2002) and central to this work:

$$\mathcal{D}_{\succcurlyeq} = \{P \in \Delta \mid \int u(f)dP \geq V(f) \text{ for all } f \in \mathcal{F}\}. \tag{1}$$

The set $\mathcal{D}_{\succcurlyeq}$ corresponds to the probability measures on S that induce subjective expected utility (SEU) preferences that assign weakly higher expected utility to all acts.

The *core* of a capacity $\nu : \Sigma \to [0,1]$ is the set of probability distributions over S that pointwise dominate ν:

$$Core(\nu) = \{P \in \Delta \mid P(E) \geq \nu(E) \ \text{ for all } E \in \Sigma\}.$$

It is well-known that the $Core(\nu) = \mathcal{D}_{\succcurlyeq}$ within the CEU model (see Ghirardato and Marinacci, 2002, Corollary 13).

2.2. The definitions of ambiguity aversion by Schmeidler as well as Ghirardato and Marinacci

As outlined in the introduction, we show that the extreme levels of our hierarchy introduced below correspond to two of the most prominent definitions of ambiguity aversion. The first definition is suggested by Schmeidler (1989) and states that the DM always has a preference for mixing.[5]

Axiom 7 (Uncertainty Aversion)**.** *If $f, g \in \mathcal{F}$ and $\alpha \in [0,1]$, then $f \sim g$ implies $\alpha f + (1-\alpha)g \succcurlyeq g$.*

[5] Ghirardato and Marinacci (2002) call the axiom "ambiguity hedging". In this paper we stick to the original name proposed by Schmeidler (1989).

Schmeidler suggests that this axiom characterizes ambiguity aversion as mixing smooths out utility distributions across states and thus reduces ambiguity. An ambiguity averse DM is therefore better off.

Definition 1 (Schmeidler-Ambiguity-Aversion)**.** A preference relation $\succcurlyeq$ is Schmeidler-ambiguity-averse if it satisfies Axiom 7.

The second definition is suggested by Ghirardato and Marinacci (2002). The intuition of their definition is that a preference relation reveals ambiguity aversion if it is more ambiguity averse then some SEU preference, which Ghirardato and Marinacci suggest are ambiguity neutral.

Definition 2 (GM-Ambiguity-Aversion)**.** A preference relation $\succcurlyeq$ is GM-ambiguity-averse if there exists an SEU preference $\succcurlyeq_{SEU}$ such that for all $f \in \mathcal{F}$ and $x \in X$:

$$f \succcurlyeq x \text{ implies } f \succcurlyeq_{SEU} x.$$

Ghirardato and Marinacci (2002) show that an IB preference $\succcurlyeq$ is ambiguity averse according to their definition if and only if $\mathcal{D}_{\succcurlyeq} \neq \emptyset$.

It is well-known that within the class of IB preferences, these two definitions have a logical relationship:[6]

$$\succcurlyeq \text{ is Schmeidler-ambiguity-averse} \implies \succcurlyeq \text{ is GM-ambiguity-averse.}$$

Schmeidlers definition is thus stronger than the one of Ghirardato and Marinacci. Outside the class of IB preferences this relationship breaks down, see Example 2 in Cerreia-Vioglio et al. (2011) for a demonstration.

3. The Hierarchy of Ambiguity Aversion

For $k \in \{1, \dots, |S|\}$, the following axiom states a preference for mixing among acts if the mixture constitutes a k-act.[7]

Axiom 8 (k - Ambiguity - Aversion)**.** *If* $f_1, \dots, f_n \in \mathcal{F}$, $\alpha_1, \dots, \alpha_n > 0$, $\alpha_1 + \cdots + \alpha_n = 1$, *and* $\alpha_1 f_1 + \cdots + \alpha_n f_n = f \in \mathcal{F}_k$, *then* $f_1 \sim \cdots \sim f_n$ *implies* $f \succcurlyeq f_1$.

[6] See section 5 of Cerreia-Vioglio et al. (2011).

[7] Chateauneuf and Tallon (2002) call 1-ambiguity-aversion "sure diversification", Chandrasekher et al. (2021) call it "preference for sure diversification".

The following definition suggests itself.

Definition 3 (k-Ambiguity-Aversion)**.** A preference relation reveals k-ambiguity-aversion if it satisfies the axiom k-ambiguity-aversion.

A DM who is k-ambiguity-averse is guaranteed to have a preference for mixing if the mix results in an act with at most k different outcomes. Consequently, the strength of k-ambiguity-aversion increases with k, justifying the term *hierarchy*.

We furthermore suggest that the different levels of the hierarchy represent increasingly strong levels of ambiguity aversion. As pointed out in Schmeidler (1989), mixing among acts smooths out utility distributions, a phenomenon that is intuitively linked to ambiguity aversion. A 1-ambiguity-averse DM is guaranteed to prefer a mix if it constitutes a constant act, which ensures that all ambiguity is eliminated. No other restrictions on the DM's preference for mixing are being made. A 2-ambiguity-averse DM is also guaranteed to prefer mixes that constitute a constant act but in addition is also guaranteed to prefer mixes that constitute a binary act. The 1-ambiguity-averse DM may not prefer such a mix since the smoothing out of utility distributions constituting a binary act are not necessarily sufficiently compelling. The 1-ambiguity-averse DM can thus be classified as less ambiguity averse than the 2-ambiguity-averse DM. Similar reasoning applies for all other values of k: As k increases, the preference for mixing among acts becomes more pronounced and thus reflects higher levels of ambiguity aversion.

The following example shows that the different levels of our hierarchy are indeed meaningful. That is for all $k \in \{1, \ldots, |S| - 1\}$ there exist IB preferences which are k-ambiguity-averse but not $(k+1)$-ambiguity-averse. The constructed preference relations have a *vector expected utility* representation (Sinischalchi, 2009). Thus a special case of this influential model serves as an example for the different levels of our hierarchy.

Example 1. For simplicity assume that acts map to utilities. Let P be the uniform distribution on S. For an act $f \in \mathcal{F}$, let $x_1(f) \leq x_m(f)$ be the consequences of f in ascending order. For $k \in \{1, \ldots, |S| - 1\}$, let $\Delta_k(f)$ be the (weakly) k-biggest utility difference amongst the neighboring consequences in the ascending order of consequences of f, e.g. if f maps to the utilities $2, 5, 6$ and 10, then $\Delta_1(f) = 4$, $\Delta_2(f) = 3$ and $\Delta_3(f) = 1$.

Consider the preference relation $\succcurlyeq$ induced by the functional

$$V(f) = \int f dP - \frac{\Delta_k(f)}{|S|}.$$

It is easily checked that $\succsim$ is an IB preference.[8] We now show that $\succsim$ satisfies k-ambiguity-aversion, but not $(k+1)$-ambiguity-aversion. The key idea here is that all k-acts are evaluated at the uniform distribution P, but all other acts are "penalized" by $\frac{\Delta_k(f)}{|S|}$ (in a way that monotonicity is not violated). Clearly $\mathcal{D}_{\succsim} = \{P\}$.[9]

Consider $f, f_1, \ldots, f_n \in \mathcal{F}$ such that $f_1 \sim \cdots \sim f_n$ and $\sum_{i=1}^{n} \alpha_i f_i = f$. If $f \in \mathcal{F}_k$ then $\Delta_k(f) = 0$ and therefore $V(f) = \int f dP = \sum_{i=1}^{n} \alpha_i E_P(f_i) \geq \sum_{i=1}^{n} \alpha_i V(f_i) = V(f_1)$, thus k-ambiguity-aversion holds.

To see that $(k+1)$-ambiguity-aversion fails consider the acts f_1 and f_2 defined by $f_1(s_i) = 2i\mathbb{1}_{i\in\{1,\ldots,k-1\}}$ and

$$f_2(s_i) = \begin{cases} 4, & \text{for } i \in \{1, 2\} \\ 2i, & \text{for } i \in \{3, \ldots, k-1\} \\ -2, & \text{for } i = k \\ 0, & \text{for } i \in \{k+1, \ldots, |S|\} \end{cases}.$$

Clearly $f_1, f_2 \in \mathcal{F}_k$. The k-biggest utility difference of these two acts is 0, thus $V(f_1) = V(f_2) = 2\frac{\sum_{i=1}^{k-1} i}{|S|}$, in particular $f_1 \sim f_2$. Consider the act $f = \frac{1}{2}f_1 + \frac{1}{2}f_2$ which maps to $k+1$ different utilities, thus $f \in \mathcal{F}_{k+1}$. The expectation of f given P is also $\frac{\sum_{i=1}^{k-1} i}{|S|}$ and the k-biggest utility difference of f is 1, thus $V(f) = \frac{\sum_{i=1}^{k-1} i}{|S|} - \frac{1}{|S|} < V(f_1)$.

[8] Among the five axioms, monotonicity is the only tricky one. Due to Axiom 1 it suffices to consider acts that differ only in one state. Let $f, g \in \mathcal{F}$ be identical except in $s' \in S$ where $f(s') > g(s')$. The absolute difference between the k-biggest differences of the two acts can be at most $f(s') - g(s')$. Thus $V(f) - V(g) \geq \frac{f(s')-g(s')}{|S|} - \frac{f(s')-g(s')}{|S|} = 0$.

[9] It is obvious that $V(f) \leq \int f dP$ for all $f \in \mathcal{F}$ since $\frac{\Delta_k(f)}{|S|}$ is non-positive. Furthermore for any prior $Q \in \Delta$ with $Q \neq P$ we can construct an act $f \in \mathcal{F}$ such that $\int f dQ < V(f)$, thus $Q \notin \mathcal{D}_{\succsim}$.

4. Results

4.1. Main theorem

Our main theorem shows that for each level of the hierarchy there is a specific relationship to the set of dominating measures. The proof is inspired by Schmeidler (1972). Theorem 2.2 of this paper on cooperative game theory provides insights on specific set-functions. Although there is at first no obvious relationship between the results, our proof builds on the discovery that Schmeidler's insights can be generalized to the class of representation functionals of IB preferences, see Lemma 3 in the appendix.

Driven by this result, we show that the CEU model is endowed with a specific hierarchical structure (see section 4.2).

Theorem 1. *Let $\succcurlyeq$ be an IB preference on $\mathcal{F}$. Let $V : \mathcal{F} \to \mathbb{R}$ and $u : X \to \mathbb{R}$ be a representation functional and utility function of $\succcurlyeq$. The following are equivalent:*

(i) $\succcurlyeq$ *is k-ambiguity-averse.*

(ii) $V(f) = \min_{P \in \mathcal{D}_{\succcurlyeq}} \int u(f)dP$ *for all* $f \in \mathcal{F}_k$.

The theorem shows that a k-ambiguity-averse DM evaluates each k-act by the minimum expected utility with respect to the set of dominating measures. Thus all k-acts are evaluated at the best possible scenario, given the nature of the set $\mathcal{D}_{\succcurlyeq}$.[10] Acts with more than k outcomes may be evaluated worse than the minimum expected utility with respect to $\mathcal{D}_{\succcurlyeq}$.

The following corollary shows that a DM is 1-ambiguity-averse if and only if she is GM-ambiguity-averse and $|S|$-ambiguity-averse if and only if she is Schmeidler-ambiguity-averse. Our hierarchy is thus a unifying framework for these two influential definitions which conform to its extremes. This provides further justification to interpret Axiom 8 as an axiomatic hierarchy of ambiguity aversion.

Corollary 1. *Let $\succcurlyeq$ be an IB preference. Then*

1. $\succcurlyeq$ *is GM-ambiguity-averse if and only if it is 1-ambiguity-averse.*

[10] Note that "best" is not a typo. Given the definition of $\mathcal{D}_{\succcurlyeq}$, the "best" that can happen to an act is that it is evaluated at the worst prior within the set $\mathcal{D}_{\succcurlyeq}$.

2. *$\succcurlyeq$ is Schmeidler-ambiguity-averse if and only if it is $|S|$-ambiguity-averse.*

Part 1. follows from results in Ghirardato and Marinacci (2002) and Grant and Polak (2013), see the appendix for details. Note that 1. also implies that 1-ambiguity-aversion is equivalent to the nonemptyness of $\mathcal{D}_{\succcurlyeq}$. The proof of part 2. needs only Axiom 1, thus this equivalence holds for any weak order.

4.2. The hierarchy with Choquet expected utility

In this section we show that within the CEU model our hierarchy only has three levels, independent of the cardinality of the state space. These three levels are represented by capacities with a non-empty core (*balanced*), capacities whose values are equal to the lower envelope of the core (*exact*), and supermodular capacities (*convex*).

Definition 4. A capacity $\nu : \Sigma \to [0, 1]$ is called

(i) *Balanced* if $Core(\nu) \neq \varnothing$.

(ii) *Exact* if $\nu(E) = \min_{P \in Core(\nu)} P(E)$ for all $E \in \Sigma$.

(iii) *Convex* if $\nu(E_1) + \nu(E_2) \leq \nu(E_1 \cup E_2) + \nu(E_1 \cap E_2)$ for all $E_1, E_2 \in \Sigma$.

It is well-known that convex capacities are exact and exact capacities are balanced and that the reverse directions do not hold (see, e.g., Shapley, 1965).

Theorem 2. *Let $\succcurlyeq$ be a CEU preference relation. Then*

(i) 1-ambiguity-aversion characterizes balanced capacities.[11]

(ii) 2-ambiguity-aversion characterizes exact capacities.

(iii) k-ambiguity-aversion characterizes convex capacities for all $k \in \{3, \ldots, |S|\}$.

[11] This generalizes a known result, see Theorem 2 in Chateauneuf and Tallon (2002).

Parts (i) and (ii) follow directly from Theorem 1 since $Core(\nu) = \mathcal{D}_{\succsim}$ in the CEU model. Part (ii) states that the capacity is exact whenever a CEU DM is 2-ambiguity-averse. This is, to the best of our knowledge, the first axiomatization of exact capacities. Part (iii) shows that the CEU model is incapable of distinguishing between more than three levels of the hierarchy. For all $k \geq 3$ our k-ambiguity-aversion axiom characterizes convex capacities. For $k = |S|$ this is obvious in light of Corollary 1 and Schmeidler (1989), who shows that the axiom Uncertainty Aversion characterizes convex capacities. For $k \in \{3, \ldots, |S| - 1\}$ however this is a new result. This insight suggests that Comonotonic Independence is sufficiently strong to prevent a distinction between these levels of ambiguity aversion. Certainty Independence on the other hand is weak enough to allow a proper distinction between all levels of the hierarchy as demonstrated in Example 1.

5. Conclusion

This paper introduces a hierarchical definition of ambiguity aversion: Higher levels are characterized by a stronger preference for mixing among acts. The structured set of testable axioms specified through our hierarchy may help to better understand and empirically study ambiguity aversion. We focus on ambiguity aversion, however the ambiguity loving counterparts of our definitions and theorems can be easily obtained by reversing inequalities and preferences.

We show that the extremes of our hierarchy correspond to two of the most influential definitions of ambiguity aversion: The weakest level is equivalent to the definition of Ghirardato and Marinacci (2002) and the strongest level to the definition of Schmeidler (1989).

Furthermore, the hierarchy has only three levels within the Choquet expected utility model (Schmeidler, 1989). We show that the second level characterizes exact capacities (Shapley, 1965) which provides the first axiomatic characterization of this class of preferences.

6. Proofs

Denote by B_0 the set of real-valued functions on S. If $f \in \mathcal{F}$ and $u : X \to \mathbb{R}$, $u(f)$ is the element of B_0 defined by $u(f)(s) = u(f(s))$ for all $s \in S$. For $c \in \mathbb{R}$, let $c^* \in B_0$ be the constant function taking value c. With abuse of notation the set of constant functions in B_0 is referred to as $\mathbb{R}$. The vector

$(1,\ldots,1) \in B_0$ is denoted by $\mathbb{I}_S$. If for $a,b \in B_0$ we have $a(s) \geq b(s)$ for all $s \in S$ we write $a \geq b$.

Denote by $B_{\geq 0}$ the set of elements on B_0 with only non-negative entries. For $k \in \mathbb{N}$, B_0^k denotes the set of elements of B_0 with at most k distinct outcomes. Furthermore we define $B_{\geq 0}^k := B_{\geq 0} \cap B_0^k$.

Consider a functional $I : B_0 \to \mathbb{R}$. We denote by $I_{\geq 0}$ the restriction of I to $B_{\geq 0}$. The following properties are well-known. A functional is called *C-independent* if $I(a + c^*) = I(a) + I(c^*)$ for all $a \in B_0$ and $c^* \in \mathbb{R}$. It is called *homogenous* if $I(\gamma a) = \gamma I(a)$ for all $a \in B_0$ and $\gamma \geq 0$. It is called *monotonic* if $I(a) \geq I(b)$ whenever $a \geq b$. It is called *superadditive* if $I(a+b) \geq I(a)+I(b)$ for all $a,b \in B_0$. It is called *normalized* if for all $k \in \mathbb{R}$, $I(k^*) = k$.

We call $I : B_0 \to \mathbb{R}$ *k-ambiguity averse* if for all $a_1,\ldots,a_m \in B_0$, $\alpha_1,\ldots,\alpha_m \in \mathbb{R}_{\geq 0}$ such that $\sum_{i=1}^m \alpha_i = 1$ and $\sum_{i=1}^m \alpha_i a_i \in B_0^k$ we have that $I(a_1) = ... = I(a_m)$ implies $I(\sum_{i=1}^m \alpha_i a_i) \geq I(a_1)$.

The following lemma is from Ghirardato et al. (2004) (GMM) and precisely characterizes the representation functional of IB preferences.

Lemma 1 (GMM)**.** *A binary relation $\succcurlyeq$ on $\mathcal{F}$ is an IB preference, i.e. satisfies Axiom 1-5, if and only if there exist a monotonic, C-independent and homogenous functional $I : B_0 \to \mathbb{R}$ and a nonconstant affine function $u : X \to \mathbb{R}$ such that*

$$f \succcurlyeq g \iff I(u(f)) \geq I(u(g)).$$

Moreover, I is unique and u is unique up to positive affine transformations.

In light of this result we call a functional $I : B_0 \to \mathbb{R}$ an *IB functional* if it is monotonic, C-independent and homogenous. Proposition 1 in Cerreia-Vioglio et al. (2011) shows that I is normalized, in particular $I(0^*) = 0$ and $I(\mathbb{I}_S) = 1$.

The set of bounded, finitely additive set functions on Σ is denoted by $ba(\Sigma)$. The set of probabilities on Σ is denoted by $pc(\Sigma)$. For $I : B_0 \to \mathbb{R}$ define

$$\mathcal{E}_I = \{P \in pc(\Sigma) | \sum_{s\in S} a(s)P(s) \geq I(a) \text{ for all } a \in B_0\}.$$

The following two functions are needed for the proof of Theorem 1.

Definition 5. For an IB functional $I : B_0 \to \mathbb{R}$ with $\mathcal{E}_I \neq \emptyset$ define the following functions $\hat{V} : B_0 \to \mathbb{R}$ and $\tilde{V} : B_0 \to \mathbb{R}$:

$$\hat{V}(a) := \inf \left\{ \sum_S P(s)a(s) | P \in \mathcal{E}_I \right\}.$$

$$\tilde{V}(a) :=$$
$$\sup \left\{ \sum \alpha_i I(a_i) - \beta | (\alpha_i, a_i) \text{ finite sequence in } \mathbb{R}_{\geq 0} \times B_0, \beta \in \mathbb{N}, \sum \alpha_i a_i - \beta \mathbb{I}_S \leq a \right\}.$$

First we prove some properties of $\tilde{V}$ which we need later.

Lemma 2. *$\tilde{V}$ is monotonic, homogenous, superadditive, C-independent and normalized.*

Proof. Superadditivity can be shown following Schmeidler (1972): Let $a, b \in B_0$ and $\varepsilon > 0$. Then, there are finite sequences (α_i, a_i) and (α_i', b_i) in $\mathbb{R}_{\geq 0} \times B_0$ and $\beta, \beta' \in \mathbb{N}$ such that

$$\sum \alpha_i a_i - \beta \mathbb{I}_S \leq a, \tag{2}$$

$$\sum \alpha_i' b_i - \beta' \mathbb{I}_S \leq b, \tag{3}$$

$$\sum \alpha_i I(a_i) - \beta \geq \tilde{V}(a) - \varepsilon/2 \tag{4}$$

and

$$\sum \alpha_i' I(b_i) - \beta' \geq \tilde{V}(b) - \varepsilon/2. \tag{5}$$

The two inequalities (2) and (3) imply

$$\sum \alpha_i a_i + \sum \alpha_i' b_i - (\beta + \beta')\mathbb{I}_S \leq a + b.$$

This together with (4) and (5) implies

$$\begin{aligned} \tilde{V}(a+b) &\geq \sum \alpha_i I(a_i) + \sum \alpha_i' I(b_i) - (\beta + \beta') \\ &\geq \tilde{V}(a) + \tilde{V}(b) - \varepsilon. \end{aligned}$$

Since ϵ was arbitrary we have the required $\tilde{V}(a+b) \geq \tilde{V}(a) + \tilde{V}(b)$.

We now show $\tilde{V}(0^*) = 0$. Monotonicity of I implies that $I(a) \geq 0$ for all $a \geq 0^*$. From the definition of $\tilde{V}$ it follows that $\tilde{V}(a) \geq I(a)$ for all $a \in B_0$ which implies $\tilde{V}(a) \geq 0$ for all $a \geq 0^*$. In particular $\tilde{V}(0^*) \geq 0$. For the other direction use superadditivity: $\tilde{V}(0^*) = \tilde{V}(0^* + 0^*) \geq \tilde{V}(0^*) + V(0^*)$ which implies $\tilde{V}(0^*) \leq 0$.

For monotonicity, assume that $a \geq b$. Superadditivity implies $\tilde{V}(a) \geq \tilde{V}(b) + \tilde{V}(a-b)$. Since $a - b \geq 0^*$ we have $\tilde{V}(a-b) \geq 0$ which implies the required $\tilde{V}(a) \geq \tilde{V}(b)$.

For homogeneity, consider $a \in B_0$ and $\gamma \geq 0$. Clearly $\tilde{V}(\gamma a) = \gamma \tilde{V}(a)$ when $\gamma = 0$ since $\tilde{V}(0^*) = 0$. So assume $\gamma > 0$. Consider a finite sequence (α_i, a_i) in $\mathbb{R}_{\geq 0} \times B_0$, $\beta \in \mathbb{N}$. Assume that $\sum \alpha_i a_i - \beta \mathbb{I}_S \leq a$, which implies $\sum \alpha_i I(a_i) - \beta \leq \tilde{V}(a)$. Of course also $\sum \alpha_i (\gamma a_i) - \gamma\beta \mathbb{I}_S \leq \gamma a$, which implies $\sum \alpha_i I(\gamma a_i) - \gamma \beta \leq \tilde{V}(\gamma a)$. Homogeneity of I implies that $\gamma(\sum \alpha_i I(a_i) - \beta \mathbb{I}_S) \leq \tilde{V}(\gamma a)$. This implies $\gamma \tilde{V}(a) \leq \tilde{V}(\gamma a)$.

For the other direction assume that $\sum \alpha_i a_i - \beta \mathbb{I}_S \leq \gamma a$, which implies $\sum \alpha_i I(a_i) - \beta \leq \tilde{V}(\gamma a)$. Of course also $\sum \alpha_i \frac{a_i}{\gamma} - \frac{\beta}{\gamma} \mathbb{I}_S \leq a$, which implies $\sum \alpha_i I(\frac{a_i}{\gamma}) - \frac{\beta}{\gamma} \leq \tilde{V}(a)$. Homogeneity of I implies that $\frac{1}{\gamma}(\sum \alpha_i I(a_i) - \beta \mathbb{I}_S) \leq \tilde{V}(a)$. This implies $\tilde{V}(\gamma a) \leq \gamma \tilde{V}(a)$. Overall we thus have the required $\tilde{V}(\gamma a) = \gamma \tilde{V}(a)$.

For C-independence consider $a \in B_0$ and $c \in \mathbb{R}$. We need to show $\tilde{V}(a + c^*) = \tilde{V}(a) + \tilde{V}(c^*)$. The $\geq$ direction follows from superadditivity of $\tilde{V}$. For the $\leq$ direction let

$$\sum \alpha_i a_i - \beta \mathbb{I}_S \leq a + c\mathbb{I}_S.$$

Rearrange to get

$$\sum \alpha_i a_i - (\beta + c)\mathbb{I}_S \leq a.$$

From the definition of $\tilde{V}$ we get

$$\sum \alpha_i I(a_i) - (\beta + c) \leq \tilde{V}(a),$$

implying

$$\sum \alpha_i I(a_i) - \beta \leq \tilde{V}(a) + c.$$

Since a_i, α_i, β were arbitrary and $c = I(c^*) \leq \tilde{V}(c^*)$ it holds that

$$\tilde{V}(a + c^*) \leq \tilde{V}(a) + \tilde{V}(c^*).$$

We now show $\tilde{V}(\mathbb{I}_S) = 1$. Consider $\sum \alpha_i a_i - \beta \mathbb{I}_S \le \mathbb{I}_S$. We have for all $P \in pc(\Sigma)$, and in particular for those in the non-empty $\mathcal{E}_I$, that $\sum \alpha_i \sum_S P(s) a_i(s) - \beta \le 1$. Thus for $P \in \mathcal{E}_I$ we have $\sum \alpha_i I(a_i) - \beta \le \sum \alpha_i \sum_S P(s) a_i(s) - \beta \le 1$. This implies $\tilde{V}(\mathbb{I}_S) \le 1$. The $\ge$ direction holds since $\tilde{V}(\mathbb{I}_S) \ge I(\mathbb{I}_S) = 1$. The rest of normalization follows from homogeneity. □

The following Lemma establishes the remarkable observation that $\hat{V}$ and $\tilde{V}$ are identical. This result is crucial for the proof of our main theorem. The proof is inspired by Schmeidler (1972) where a similar approach is used.

Lemma 3. *Let $I : B_0 \to \mathbb{R}$ be an IB functional with $\mathcal{E}_I \neq \emptyset$. Then for all $b \in B_{\ge 0}$*

$$\hat{V}(b) = \tilde{V}(b).$$

The lemma restricts attention to $B_{\ge 0}$. The extension to B_0 is straightforward due to C-independence of $\tilde{V}$ and $\hat{V}$ (C-independence of $\hat{V}$ not shown but obvious). Indeed throughout the proof of Lemma 3 we often limit attention to $B_{\ge 0}$ and tacitly use C-independence to generalize insights to B_0.

To prove the Lemma a separating hyperplane theorem is used.

Theorem 3 (Separating Hyperplane Theorem, Dunford and Schwartz,1958)**.** *In a linear topological space, any two disjoint convex sets, one of which has an interior point, can be separated by a non-zero continuous linear functional.*

Proof of Lemma 3. $\underline{\hat{V} \ge \tilde{V}}$: Consider $b \in B_{\ge 0}$ and an arbitrary $P \in \mathcal{E}_I$. Assume that

$$\sum_{i=1}^{n} \alpha_i a_i - \beta \mathbb{I}_S \le b.$$

It follows that

$$\begin{aligned}\sum_S P(s) b(s) &\ge \sum_S P(s) \left(\sum_{i=1}^{n} \alpha_i a_i(s) - \beta \right) \\ &= \sum_{i=1}^{n} \alpha_i \sum_S P(s) a_i(s) - \beta \\ &\ge \sum_{i=1}^{n} \alpha_i I(a_i) - \beta.\end{aligned}$$

It follows that $\sum\limits_{S} P(s)b(s) \geq \tilde{V}(b)$. Since P was arbitrary and $\mathcal{E}_I$ is compact, a fact noted in Ghirardato and Marinacci (2002), we have that $\hat{V}(b) \geq \tilde{V}(b)$.

$\underline{\hat{V} \leq \tilde{V}}$:

Consider $b \in B_{\geq 0}$. The idea is to construct a set function $F \in pc(\Sigma)$ which is an element of $\mathcal{E}_I$ such that $F(b) = \tilde{V}(b)$, which implies $\tilde{V}(b) \geq \hat{V}(b)$.

Case 1: $\tilde{V}(b) \neq 0$. Note that $b \in B_{\geq 0}$, so monotonicity of $\tilde{V}$ implies $\tilde{V}(b) > 0$. Consider the following sets

$$A = \{a \in B_{\geq 0} | \tilde{V}(a) \geq 1\}, \qquad B = \{a \in B_{\geq 0} | a \leq \mathbb{I}_S\},$$
$$C = \left\{ a \in B_{\geq 0} | a \leq \frac{b}{\tilde{V}(b)} \right\}.$$

The set A is convex due to the superadditivity of $\tilde{V}$ and the set $D := Co(B \cup C)$ is convex by construction. The set $int(A)$, the interior of A, is nonempty. This is because there exists a $\gamma > 0$ such that $\tilde{V}(\gamma b) = \gamma \tilde{V}(b) > 1$, so $\gamma b \in int(A)$. To apply Theorem 3 for the sets $int(A)$ and D we need to show that $int(A) \cap D = \emptyset$. It suffices to show that $\tilde{V}(a) \leq 1$ for all $a \in D$. Due to monotonicity of $\tilde{V}$ it suffices to show that $\tilde{V}\left(r\mathbb{I}_S + (1-r)\frac{b}{\tilde{V}(b)}\right) = 1$ for $0 \leq r \leq 1$. It follows directly from C-independence, homogeneity and normalization.

The requirements for Theorem 3 for the sets $int(A)$ and D are therefore fulfilled and we can conclude the existence of a linear functional $F : B_{\geq 0} \to \mathbb{R}$ such that

$$F(a) \leq F(\mathbb{I}_S) = 1 = F\left(\frac{b}{\tilde{V}(b)}\right) \leq F(c)$$

for all $a \in D$ and $c \in A$. The equalities are due to the fact that both $\mathbb{I}_S$ and $\frac{b}{\tilde{V}(b)}$ are in $A \cap D$. Since F is linear and $F(\mathbb{I}_S) = 1$ we have that $F \in pc(\Sigma)$. The second equality and the homogeneity of F furthermore imply that $F(b) = \tilde{V}(b)$.

We need to show that $\tilde{V}(a) \leq F(a)$ for all $a \in B_{\geq 0}$, which implies that $F \geq \tilde{V}$ and so $F \in \mathcal{E}_I$.

Fix some $a \in B_{\geq 0}$. The linearity of F implies that $F(a) \geq 0$. If $\tilde{V}(a) = 0$ then $F(a) \geq \tilde{V}(a)$. So assume that $\tilde{V}(a) > 0$. There exists an $r > 0$ such that $\tilde{V}(ra) = 1$. This implies that $ra \in A$ and therefore $F(ra) \geq 1$. Homogeneity of $\tilde{V}$ and F then implies the required $\tilde{V}(a) \leq F(a)$. Therefore

$$\tilde{V}(b) = F(b) \geq \inf \left\{ \sum_S P(s)b(s) | P \in \mathcal{E}_I \right\} = \hat{V}(b).$$

<u>Case 2:</u> $\tilde{V}(b) = 0$. Define A and B as above and

$$C = \bigcup_{n=1}^{\infty} \{a \in B_{\geq 0} | a \leq nb\}.$$

To show $int(A) \cap D = \emptyset$ we conclude with C-linearity, homogeneity and normalization that for $0 \leq r \leq 1$

$$\begin{aligned} \tilde{V}(r\mathbb{I}_S + (1-r)nb) &= r + (1-r)n\tilde{V}(b) \\ &= r \leq 1. \end{aligned}$$

So again we can use Theorem 3 to get the separating function F with $F(\mathbb{I}_S) = 1$, which implies $F \in pc(\Sigma)$. We need to show $F(b) = 0$. Assume that $F(b) > 0$. Then there exists an n such that $F(nb) > 1$. So $nb \in A$ as well as $nb \in D$ which contradicts the separating property of F. Therefore $F(b) = 0$. To show $F(a) \geq \tilde{V}(a)$ for all $a \in I_{\geq 0}$ the above approach is used. So $F \in pc(\Sigma)$ which implies $\tilde{V}(b) \geq \hat{V}(b)$. This finishes the proof.

□

Lemma 4. *Let $\succcurlyeq$ be an IB preference and $I : B_0 \to \mathbb{R}$ its corresponding functional from Lemma 1. Then the following are equivalent:*

1. $\mathcal{D}_{\succcurlyeq} \neq \emptyset$.
2. $\mathcal{E}_I \neq \emptyset$.
3. $\succcurlyeq$ *is 1-ambiguity-averse.*
4. I *is 1-ambiguity-averse.*

Furthermore, $\succcurlyeq$ is k-ambiguity-averse if and only if I is k-ambiguity-averse.

Proof. The equivalences between 1. and 2. as well as 3. and 4. are clear. For the equivalence between 1. and 3., note that *mean-dispersion* preferences as introduced Grant and Polak (2013) have IB preferences as a special case as shown in their Corollary 3 (c). Thus Corollary 5 (a) from Grant and Polak (2013) can be used which states the required equivalence between 1. and 3.

The "furthermore" part follows from the fact that the functional property k-ambiguity-aversion is a restatement of the k-ambiguity-aversion axiom. $\square$

Since k-ambiguity-aversion increases in strength with k, Lemma 4 implies in particular that $\mathcal{D}_{\succcurlyeq}$ (and $\mathcal{E}_I$) is nonempty when $\succcurlyeq$ (I) is k-ambiguity-averse for any $k \in \{1, \dots, |S|\}$.

We can now prove Theorem 1.

Proof of Theorem 1. Let $\succcurlyeq$ be an IB preference on $\mathcal{F}$. Let $I : B_0 \to \mathbb{R}$ be the corresponding functional from Lemma 1. Fix some $k \in \{1, \dots, |S|\}$.

"1. $\implies$ 2.": Lemma 4 guarantees that $\mathcal{D}_{\succcurlyeq} \neq \emptyset$ and thus $\mathcal{E}_I \neq \emptyset$. Due to the compactness of $\mathcal{D}_{\succcurlyeq}$ (and $\mathcal{E}_I$), *inf* and *min* can be used interchangably, i.e. the infimum expectation of some act $f \in \mathcal{F}$ (or vector $a \in B_0$) over $\mathcal{D}_{\succcurlyeq}$ (or $\mathcal{E}_I$) is equal to the well-defined minimum expectation over the respective set. Assume that 2. fails. This implies that there exists an $a \in B^k_{\geq 0}$ such that $I(a) < \min_{P \in \mathcal{E}_I} \sum_S a dP = \hat{V}(a)$. Lemma 3 implies that $I(a) < \tilde{V}(a)$. It follows that there exist $(\alpha_i, a_i)_{i=1}^m \in \mathbb{R}^m_+ \times (B_0)^m, \beta \in \mathbb{N}$ with

$$\sum_{i=1}^m \alpha_i a_i - \beta \mathbb{I}_S = a \tag{6}$$

such that

$$I(a) < \sum_{i=1}^m \alpha_i I(a_i) - \beta. \tag{7}$$

We get strict equality in (6) since we can always add other non-negative terms of the form $\alpha_i a_i$ and since monotonicity holds.

Define $b_i = \alpha_i a_i$ for $i \in \{1, \dots m\}$ and $b_{m+1} = -\beta$. By (6) it holds that

$$a = \sum_{i=1}^{m+1} b_i.$$

Due to C-independence of I we can find constants $c_1, \dots, c_{m+1} \in \mathbb{R}$ with $\sum_{i=1}^{m+1} c_i = 0$ such that for $b'_i := b_i + c_i, i \in \{1, \dots, m+1\}$ it holds that

$$I(b'_1) = \dots = I(b'_{m+1}).$$

Obviously $a = \sum_{i=1}^{m+1} b_i'$. For $b_i^\dagger := (m+1)b_i'$ it holds that

$$I(b_1^\dagger) = \cdots = I(b_{m+1}^\dagger) \tag{8}$$

and

$$a = \sum_{i=1}^{m+1} \frac{1}{m+1} b_i^\dagger. \tag{9}$$

Now

$$\begin{aligned} I(a) &\overset{(7)}{<} \sum_{i=1}^{m} \alpha_i I(a_i) - \beta \\ &= \sum_{i=1}^{m+1} I(b_i) \\ &= \sum_{i=1}^{m+1} I(b_i') \\ &= \sum_{i=1}^{m+1} \frac{1}{m+1} I(b_i^\dagger) \\ &= I(b_1^\dagger). \end{aligned}$$

Therefore $I(b_1^\dagger) > I(a)$ and this in combination with (8) and (9) shows that I violates k-ambiguity-aversion. So $\succcurlyeq$ violates k-ambiguity-aversion due to Lemma 4.

"2. $\implies$ 1.": Consider $b \in B_0^k$ and $a_1, \ldots, a_m$ such that $I(a_1) = \cdots = I(a_m)$ with $\sum_{i=1}^{m} \alpha_i a_i = b$, $\sum_{i=1}^{m} \alpha_i = 1$ and $\alpha_i \geq 0$.

Consider some $Q \in \arg\min_{P \in \mathcal{E}_I} \sum_{s \in S} b(s)P(s)$. It holds that

$$\begin{aligned}
I(a_1) &= \sum_{i=1}^{m} \alpha_i I(a_i) \\
&\leq \sum_{i=1}^{m} \alpha_i \min_{P \in \mathcal{D}} \sum_{s \in S} P(s) a_i(s) \\
&\leq \sum_{i=1}^{m} \alpha_i \sum_{s \in S} Q(s) a_i(s) \\
&= \sum_{s \in S} \sum_{i=1}^{m} \alpha_i Q(s) a_i(s) \\
&= \sum_{s \in S} Q(s) b(s) \\
&= I(b),
\end{aligned}$$

where the second inequality holds since $Q \in \mathcal{E}_I$ and the last equality holds by assumption. We have therefore shown that $I(b) \geq I(a_1)$, so I satisfies k-ambiguity-aversion. Lemma 4 implies that $\succcurlyeq$ satisfies k-ambiguity-aversion.

□

Proof of Corollary 1. For 1., Ghirardato and Marinacci (2002) show in Theorem 12 that an IB preference relation is GM-ambiguity-averse if and only if $\mathcal{D}_{\succcurlyeq} \neq \emptyset$. Corollary 5 (a) from Grant and Polak (2013) implies $\mathcal{D}_{\succcurlyeq} \neq \emptyset$ if and only if 1-ambiguity-aversion holds. Thus these two results imply the required.

The nontrivial direction of part 2. is that Schmeidler-ambiguity-aversion implies $|S|$-ambiguity-aversion. We prove it by induction over n, the number of indifferent acts. The base case is $n = 2$ which holds by assumption. Assume that it holds for $n - 1$. We need to show that it then also holds for n.

Assume that $f_1 \sim \cdots \sim f_n$ with $\alpha_i \geq 0$ for all $i \in \{1, \ldots, n\}$, $\sum_{i=1}^{n} \alpha_i = 1$ and $\sum_{i=1}^{n} \alpha_i f_i = f$. Consider the act

$$f' = \frac{\sum_{i=2}^{n} \alpha_i f_i}{1 - \alpha_1}.$$

Note that it holds that $\frac{\sum_{i=2}^{n} \alpha_i}{1-\alpha_1} = 1$. So the act f' is a convex mix of the $n-1$ acts $f_2, \ldots, f_n$. By the induction hypothesis we have $f' \succcurlyeq f_2$ which implies that $f' \succcurlyeq f_1$.

It holds that $f = \alpha_1 f_1 + (1-\alpha_1) f'$. Therefore by Schmeidler-ambiguity-aversion we have that

$$f \succcurlyeq f_1.$$

This finishes the induction step. □

Proof of Theorem 2. Let $\succcurlyeq$ be a CEU preference relation with capacity ν and let $I : B_0 \to \mathbb{R}$ and $u : X \to \mathbb{R}$ be the functionals from Lemma 1.

Corollary 13 in Ghirardato and Marinacci (2002) states that the core of a capacity always corresponds to the set $\mathcal{D}_{\succcurlyeq}$. Thus (i) and (ii) follow from Theorem 1.

For (iii) the case $|S| \leq 3$ follows from Corollary 1. Thus assume that $|S| > 3$ and that k-ambiguity-aversion holds for some $k \geq 3$. This implies $Core(\nu) \neq \emptyset$. Assume for contradiction that ν is not convex. The idea is to show that this leads to a violation of 3-ambiguity-aversion, and thus k-ambiguity-aversion.

Since ν is not convex, there exist events $E, F \in \Sigma$ such that

$$\nu(E) + \nu(F) > \nu(E \cup F) + \nu(E \cap F).$$

Consider the act f which results in x on $E \cap F$, y on $E \cup F \setminus (E \cap F)$ and z otherwise, so $f \in \mathcal{F}_3$. Assume that $u(x) = 2, u(y) = 1$ and $u(z) = 0$ (we achieve the existence of such an act through Non-degeneracy and rescaling of the utility function). The Choquet expected utility of f is

$$\begin{aligned}
\int u(f)\ d\nu &= 2\nu(E \cap F) + 1(\nu(E \cup F) - \nu(E \cap F)) \\
&= \nu(E \cap F) + \nu(E \cup F) \\
&< \nu(E) + \nu(F) \\
&\leq \min_{P \in Core(\nu)} P(E) + P(F) \\
&= \min_{P \in Core(\nu)} P(E \cup F) + P(E \cap F) \\
&\leq \min_{P \in Core(\nu)} \int f\ dP.
\end{aligned}$$

Thus we have found an $f \in \mathcal{F}_3$ such that $V(f) < \min_{P \in Core(\nu)} \int u(f) dP$.

$\square$

References

Amarante, M. (2009), 'Foundations of neo-bayesian statistics', *Journal of Economic Theory* **144**(5), 2146–2173.

Anscombe, F. and Aumann, R. (1963), 'A definition of subjective probability', *The Annals of Mathematical Statistics* **34**, 199–205.

Camerer, C. and Weber, M. (1992), 'Recent developments in modeling preferences: Uncertainty and ambiguity', *Journal of Risk and Uncertainty* **5**, 325–370.

Cerreia-Vioglio, S., Ghirardato, P., Maccheroni, F., Marinacci, M. and Siniscalchi, M. (2011), 'Rational preferences under ambiguity', *Economic Theory* **48**(2-3), 341–375.

Chandrasekher, M., Frick, M., Iijima, R. and Le Yaouanq, Y. (2021), 'Dual-self representations of ambiguity preferences', *Mimeo* .

Chateauneuf, A. and Tallon, J. (2002), 'Diversification, convex preferences and non-empty core in the choquet expected utility model', *Economic Theory* **19**, 509–523.

Choquet, G. (1954), 'Theory of capacities', *Annales de L' Institut Fourier* **5**, 131–295.

Dunford, N. and Schwartz, J. T. (1958), *Linear operators part I: general theory*, Vol. 7, Interscience publishers New York.

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

Epstein, L. (1999), 'A definition of uncertainty aversion', *Review of Economic Studies* **66**, 579–608.

Ghirardato, P., Maccheroni, F. and Marinacci, M. (2004), 'Differentiating ambiguity and ambiguity attitude', *Journal of Economic Theory* **118**, 133–173.

Ghirardato, P. and Marinacci, M. (2002), 'Ambiguity made precise: A comparative foundation', *Journal of Economic Theory* **102**, 251–289.

Grant, S. and Polak, B. (2013), 'Mean-dispersion preferences and constant absolute uncertainty aversion', *Journal of Economic Theory* **148**, 1361–1398.

Hartmann, L. (2018), Perceived Ambiguity, Ambiguity Attitude and Strategic Ambiguity in Games, PhD thesis, University of Exeter.

Hartmann, L. and Kauffeldt, T. F. (2018), 'A hierarchy of ambiguity aversion', *Working Paper* .

Klibanoff, P., Marinacci, M. and Mukerji, S. (2005), 'A smooth model of decision making under ambiguity', *Econometrica* **73**, 1849–1892.

Machina, M. and Schmeidler, D. (1992), 'A more robust definition of subjective probability', *Econometrica* **60**, 745–780.

Savage, L. (1954), *The foundations of statistics*, John Wiley and Sons, New York.

Schmeidler, D. (1972), 'Cores of exact games', *Journal of mathematical analysis and applications* **40**, 214–225.

Schmeidler, D. (1989), 'Subjective probability and expected utility without additivity', *Econometrica* **57**, 571–587.

Shapley, L. (1965), Notes on n-person games VII. cores of convex games, Technical report, Rand Corp Santa Monica California.

Sinischalchi, M. (2009), 'Vector expected utility and attitudes towards variation', *Econometrica* **77**, 801–855.

Yaari, M. (1987), 'The dual theory of choice under risk', *Econometrica* **55**, 95–115.