2-2020

Preferences with changing ambiguity aversion

Jingyi XUE
Singapore Management University, JYXUE@smu.edu.sg

Citation

XUE, Jingyi. Preferences with changing ambiguity aversion. (2020). *Economic Theory*. 69, (1), 1-60.
Available at: https://ink.library.smu.edu.sg/soe_research/2237

Economic Theory
https://doi.org/10.1007/s00199-018-1156-2

Preferences with changing ambiguity aversion

Jingyi Xue[1]

Received: 17 March 2016 / Accepted: 3 October 2018

Abstract
We provide two extensions of Gilboa and Schmeidler (J Math Econ 18:141–153, 1989)'s maxmin expected utility decision rule to accommodate a decision maker's changing ambiguity attitudes. The two rules are, respectively, a weighted maxmin rule and a variant constraint rule. The former evaluates an act by a weighted average of its worst and best possible expected utilities over a set of priors, with the weights depending on the act. The latter evaluates an act by its worst expected utility over a neighborhood of a set of approximating priors, with the size of the neighborhood depending on the act. Canonical representations of the two rules are provided for classes of preference relations that exhibit, respectively, ambiguity aversion à la Schmeidler (Econometrica 57:571–587, 1989) and ambiguity aversion à la Ghirardato and Marinacci (J Econ Theory 102:251–289, 2002). In the second part of this paper, we study wealth effect under ambiguity. We propose axioms on absolute and relative ambiguity aversion and derive three representations for the ambiguity averse preference relations exhibiting decreasing (increasing) absolute ambiguity aversion. In particular, decreasing absolute ambiguity aversion implies that as the baseline utility of an act increases, a weighted maxmin decision maker puts less weight on the worst case, and a variant constraint decision maker considers a smaller neighborhood of approximating priors.

Keywords Ambiguity · Ambiguity averse preferences · Weighted maxmin representation · Variant constraint representation · Decreasing absolute ambiguity aversion · Increasing relative ambiguity aversion · Wealth effect

This paper subsumes the previous version "Three representations of preferences with decreasing absolute uncertainty aversion (Xue 2012)", which is Chapter 2 of my PhD thesis. I appreciate Simon Grant, Atsushi Kajii, Siyang Xiong, and Hervé Moulin for invaluable discussion and suggestions. I thank the referee for constructive suggestions which have greatly improved the paper. I also thank Shurojit Chatterji, Chiaki Hara, Ehud Lehrer, Jin Li, Xin Li, Thomas Sargent, Marciano Siniscalchi, William Thomson, Stephen Wolff, and Minyan Zhu for their helpful comments. Parts of this research were carried out when I was visiting Kyoto University. I am grateful to Kyoto Institute of Economic Research for their hospitality and support.

✉ Jingyi Xue
jyxue@smu.edu.sg

1 Singapore Management University, Singapore, Singapore

JEL Classification D81

1 Introduction

A decision maker makes choices facing unknown states of the world as well as unknown probability distributions of the states. Such a situation is called Knightian uncertainty (Knight 1921) or ambiguity. It is different from a situation of risk in which the states are unknown, but the probability distribution is known. The decision maker has a preference relation over acts that yield state-contingent outcomes. A well-known decision rule axiomatized by Gilboa and Schmeidler (1989) is the maxmin expected utility (MEU) rule (see also Wald 1950a, b). A MEU decision maker evaluates an act f by

$$\min_{p \in D} E_p u(f),$$

where D is a set of priors over the states, u is a utility function over outcomes, and $E_p u(f)$ denotes the expected utility of the act f with respect to a prior p. The decision maker behaves as if he regards all priors in D as possible and pessimistically evaluates an act by its worst possible expected utility.

The MEU decision rule is a prominent rule that captures a decision maker's aversion to ambiguity. However, by considering always the worst case, it does not accommodate the possibility that a decision maker has changing ambiguity attitudes. For an example, imagine that a decision maker chooses, in the face of two unknown states, one from the following two acts that are parameterized by $t \geq 0$. The first act is ambiguous, yielding $100+t$ dollars in state 1 and t dollars in state 2. The second act is unambiguous, yielding $45+t$ dollars in both states. Assume that the decision maker is risk neutral and believes that the probability of state 1 is between 0.4 and 0.6. Thus, the worst expected payoff of the unambiguous act is $45 + t$ and that of the ambiguous act is $40 + t$. According to the MEU decision rule, regardless of the value of t, the unambiguous act is better than the ambiguous one. However, as analogous to the wealth effect under risk, people may tend to be less averse to ambiguity when the baseline payoff of an act increases. In particular, the decision maker may choose the unambiguous act when $t = 0$ and choose the ambiguous act when $t = 1000$. Similar behavioral patterns that reveal a decision maker's changing degrees of ambiguity aversion are well evidenced in the systematically designed experiments studied by Baillon and Placido (2015). The MEU decision rule does not accommodate such behavior patterns. Therefore, we consider in this paper two extensions of the MEU decision rule.

The first extension is a generalized Hurwicz rule, or a weighted maxmin (WM) decision rule. A WM decision maker evaluates an act f by

$$\lambda(u(f)) \min_{p \in D} E_p u(f) + (1 - \lambda(u(f))) \max_{p \in D} E_p u(f),$$

where D and u are the same as in the MEU rule, and λ is a function that depends on utility profiles induced by acts and takes its values in $[0, 1]$. We call D an admissible set of priors, or simply an **admissible set**, and λ a weight function. A WM decision maker behaves as if he considers not only the worst possible expected utility but also the best, and evaluates an act by a weighted average of the worst and best expected utilities. Different from the MEU rule, the weight on the worst case is not a constant equal to 1. Instead, it depends on acts. Smaller weights correspond to less ambiguity aversion. When the weight goes to 1, the decision maker becomes optimistic and thus exhibits ambiguity loving.

A WM decision rule typically has more than one representations given by different admissible sets. Indeed, for each superset of D, we can use the same u and an adjusted weight function to generate the same value for each act, so that they constitute an alternative representation of the same preference relation. Each admissible set gives rise to an upper bound and a lower bound of the value of each act. We are interested in finding the smallest admissible set since it provides the tight bounds and gives a natural normalization in view of the nesting property of the admissible sets. We call a WM representation that has the smallest admissible set a **canonical WM representation**.

Ghirardato et al. (2004) also introduce a WM decision rule with a representation that has a particular admissible set. The two representations are different in general. A preference relation that admits their WM representation may not admit a canonical WM representation. Moreover, when a preference relation admits both representations, their admissible set may not be the smallest. See Sect. 3.1.1 for examples. However, when restricted to the class of so-called ambiguity averse preference relations in the literature, the two representations coincide.

More precisely, an ambiguity averse preference relation is one that satisfies some basic axioms and exhibits ambiguity aversion à la Schmeidler (1989), which we call S-ambiguity aversion. Our first main result shows that a preference relation is an ambiguity averse preference relation if and only if it admits a canonical WM representation in which the weight function satisfies some properties. Moreover, in this case, the smallest admissible set agrees with the admissible set in the WM representation studied by Ghirardato et al. (2004).

The second extension of the MEU rule is a variant of the constraint decision rule that is introduced by Hansen and Sargent (2001) as one of their two robust decision rules. A variant constraint (VC) decision maker evaluates an act f by

$$\min_{p\in\Delta:d(p,K)\leq\sigma(u(f))} E_p u(f),$$

where u and Δ are the same as before, K is a set of priors, $d(p, K)$ is the Euclidean distance between the prior p and the set K, and σ is a function that depends on utility profiles induced by acts and takes non-negative real values. We call K an essential set of priors, or simply an **essential set**, and σ a constraint function. A VC decision maker behaves as if he considers the priors in the essential set K as best approximations of the true prior, while he is also concerned with potential misspecification of approximating priors. Thus, the decision maker considers a neighborhood of the essential set and evaluates an act by its worst expected utility with respect to the priors in the neighborhood. Importantly, the size of the neighborhood, specified by σ, depends on the act in consideration. Larger values of σ correspond to less concern about misspecification of approximating priors. When σ goes to 0, the decision maker only considers approximating priors and is not at all concerned about potential misspecification. When the constraint function σ is constant, the VC decision rule reduces to a MEU decision rule.

Similar to a WM decision rule, a VC decision rule typically has more than one representations. Indeed, for each subset of K, we can use the same u and an adjusted constraint function to generate the same value for each act, so that they constitute an alternative representation of the same preference relation. In light of the nesting property of the essential sets, we are interested in finding the largest one as a normalization. We call a VC representation that has the largest essential set a **canonical VC representation**.

While our first main result provides canonical WM representations for preference relations exhibiting S-ambiguity aversion, our second main result provides canonical VC representations for those exhibiting a different type of ambiguity aversion à la Ghirardato and Marinacci (2002), which we call GM-ambiguity aversion. Based on a notion of comparative ambiguity aversion, Ghirardato and Marinacci (2002) say that a preference relation is ambiguity averse if it is more ambiguity averse than a subjective expected utility (SEU) preference relation. We show that a preference relation satisfies some basic axioms and exhibits GM-ambiguity aversion if and only if it admits a canonical VC representation in which the constraint function satisfies a monotonicity property. Moreover, we show that the largest essential set for a preference relation is the set of all subjective priors of the SEU preference relations that are less ambiguity averse than the one in consideration.

Neither the class of preference relations admitting canonical WM representations includes the class of preference relations admitting canonical VC representations, nor the reverse. Since S-ambiguity aversion does not imply GM-ambiguity aversion, by our two representation results, there are preferences admitting canonical WM representations but not canonical VC representations. Conversely, we provide an example of a preference relation that admits a canonical VC representation but not a canonical WM representation. When a preference relation admits both types of representations, we show that the largest essential set is always a subset of the smallest admissible set. Moreover, the MEU preference relations are characterized by the coincidence of the largest essential set with the smallest admissible set.

After studying two extensions of the MEU decision rule that accommodate changing ambiguity attitudes, we focus in the second part of the paper on ambiguity averse preference relations that exhibit a particular pattern of changing ambiguity aversion. More precisely, we study the wealth effect on the class of preference relations exhibiting S-ambiguity aversion. As discussed in the motivating example and as evidenced by Baillon and Placido (2015), people tend to be less averse to ambiguity when they become better off overall. To capture this behavioral pattern, we propose an axiom of decreasing absolute ambiguity aversion. It says that if an (ambiguous) act is preferred to a constant (unambiguous) act, then this is still the case after a common improvement on both acts. With the assumption that there is an outcome that both acts yield with the same probability in all states, the common improvement on acts is formulated as the improvement on that common outcome. Under some basic axioms, such an improvement on an act amounts to a uniform increase in the utilities that it generates in all states. Thus, equivalently the axiom says that if the ambiguity of an act is previously tolerable, it is even more tolerable after the ensured utility of the act increases.

Our third main result provides three representations for the subclass of ambiguity averse preference relations exhibiting decreasing absolute ambiguity aversion. The first two representations are, respectively, a WM representation and a VC representation, on which the wealth effect has straightforward implications. That is, as the ensured utility of an act increases, a WM decision maker behaves as if he becomes less pessimistic and puts a smaller weight on the worst case, and a VC decision maker behaves as if he is less concerned with prior misspecification and considers a smaller neighborhood of approximating priors. The third representation is an ambiguity averse representation introduced by Cerreia-Vioglio et al. (2011b). The wealth effect, when reflected in this representation, amounts to a monotonicity property related to the ambiguity aversion index in the representation.

Note that the part of our third main result on the WM and VC representations is not a corollary of our first two main results. Instead of searching for canonical representations, here we impose a limit condition on each representation. First, we require that a WM decision maker tends to put the whole weight on the worst case in an extremely bad situation in which the baseline utility of an act is sufficiently low and the scale of its ambiguous part is sufficiently large. Second, we require that a VC decision maker tends to consider only approximating priors in an extremely good situation in which the baseline utility of an act is sufficiently high and the scale of its ambiguous part is sufficiently small. It turns out that under the first limit condition, the admissible set in the WM representation is actually the smallest admissible set. Similarly, under the second limit condition, the essential set in the VC representation is the largest essential set. As a result, the two representations that we derive are canonical representations, although we do not require them to be canonical.

Analogous representations are obtained for the subclass of ambiguity averse preference relations exhibiting increasing and constant absolute ambiguity aversion. As a corollary, we get two alternative representations for variational preference relations studied by Maccheroni et al. (2006), since variational preference relations constitute the subclass of ambiguity averse preference relations exhibiting constant absolute ambiguity aversion. Moreover, variational preference relations also exhibit increasing relative ambiguity aversion that we define in a similar way as decreasing absolute ambiguity aversion. Increasing relative ambiguity aversion has intuitive implications on the WM and VC representations of variational preference relations. As the scale of the ambiguity of an act increases, a WM decision maker behaves as if he is more pessimistic and puts more weight on the worst case, and a VC decision maker behaves as if he is more concerned about prior misspecification and considers a larger neighborhood of approximating priors.

After discussing the related literature, the rest of the paper is organized as follows: Section 2 introduces the model and axioms. Section 3 introduces two extensions of the MEU decision rule—the WM rule and the VC rule. Characterizations of classes of preference relations that admit respective canonical representations are provided. As an application, we present the two canonical representations of multiplier preference relations introduced by Hansen and Sargent (2001). Section 4 studies wealth effect under ambiguity and proposes an axiom of decreasing absolute ambiguity aversion.

Representations are provided for the subclass of ambiguity averse preference relations exhibiting decreasing absolute ambiguity aversion. Analogous results are provided for preference relations exhibiting increasing and constant absolute ambiguity aversion. Section 5 concludes. All the proofs are given in "Appendix".

1.1 Related literature

There have been quite a few studies of different versions of the two generalized MEU decision rules in different settings, e.g., Ghirardato et al. (2004), Cerreia-Vioglio et al. (2011a), Olszewski (2007), Gajdos et al. (2008), Kopylov (2009), Chateauneuf and Faro (2009) and Hill (2013). Among them, the most closely related works are Ghirardato et al. (2004) and Hill (2013). Ghirardato et al. (2004) introduce a WM decision rule and derive a representation with the certainty independence axiom of Gilboa and Schmeidler (1989) being imposed. The same type of representation is also obtained, after dropping certainty independence, by Cerreia-Vioglio et al. (2011a). The relationship between their WM representation and ours is elaborated in detail in Sect. 3.

Hill (2013) also axiomatizes a class of preference relations that exhibit changing ambiguity aversion across acts. In his model, each act is evaluated by the worst expected utility over a set of priors that depends on the "stakes" involved in choosing the act. This class of preference relations satisfies the axioms of our Theorem 2, so it constitutes a subclass of preference relations that admits a canonical VC representation. This subclass satisfies a stronger independence axiom, a stronger monotonicity axiom and a stronger ambiguity aversion axiom. These axioms together imply that as the stakes get bigger, the decision maker evaluates an act by its worst expected utility over a larger set of priors, as if he becomes less confident.

Regarding the wealth effect under ambiguity, Cherbonniera and Gollier (2015) propose a definition of decreasing aversion under ambiguity within the smooth ambiguity model and the α-MEU model. They consider changes in monetary wealth, but their definition does not distinguish the effect on risk aversion and that on ambiguity aversion. Cerreia-Vioglio et al. (2017) propose a definition of decreasing absolute ambiguity aversion in a general setting. They also consider changes in monetary wealth, and their definition implies that a decision maker who exhibits decreasing absolute ambiguity aversion must exhibit constant absolute risk aversion. In contrast to these two works, we investigate the effect of changes in baseline utilities on ambiguity aversion, and thus impose no restrictions on a decision maker's risk attitude. Chambers et al. (2014) provide a similar definition as ours but theirs is stronger. Our axiom is closer in spirit to Klibanoff et al. (2005)'s definition, although the authors only define constant absolute ambiguity aversion. We show that in their model, our axiom is equivalent to decreasing concavity of a second-order utility index. This is consistent with the claim of Klibanoff et al. (2005) that a second-order utility index summarizes one's ambiguity attitude in the same way as a von Neumann–Morgenstern utility function summarizes one's risk attitude. Chateauneuf and Faro (2009) assume the existence of a worst outcome and propose the "worst independence" axiom. This axiom amounts to our constant relative ambiguity aversion axiom under the assumption of constant absolute ambiguity aversion, but in general it is weaker.

2 The model

Let S be a finite set of **states of the world** that contains at least two states, and Δ the probability space over S. Let X be a set of **outcomes**. Following Maccheroni et al. (2006), we assume that X is a convex subset of some vector space. For example, X is an interval of monetary payoffs, or the set of all lotteries over a set of prizes as studied by Anscombe and Aumann (1963). An **act** is a function $f : S \to X$ that yields in each state an outcome. Let $\mathcal{F}$ be the set of all acts. With a slight abuse of notation, for each $x \in X$, we denote by x the **constant act** in $\mathcal{F}$ that yields x in all states, and we identify X with the set of all constant acts. For each pair $f, g \in \mathcal{F}$ and each $\alpha \in [0, 1]$, let $\alpha f + (1 - \alpha)g$ be the mixed act that yields in each $s \in S$ the mixed outcome $\alpha f(s) + (1 - \alpha)g(s)$. Note that mixed outcomes are well defined due to the convexity of X. A decision maker's **preference relation** is a binary relation $\succsim$ on $\mathcal{F}$. As usual, let $\succ$ and $\sim$ denote, respectively, the asymmetric and symmetric parts of $\succsim$. Given a decision maker's preference relation $\succsim$, for each $f \in \mathcal{F}$, a **certainty equivalent** of f is a constant act, denoted by x_f, between which and the act f the decision maker is indifferent.

We consider the following basic axioms on a preference relation $\succsim$.

A.1. Weak order The preference relation $\succsim$ is complete and transitive.

A.2. Risk independence For all $x, y, z \in X$ and all $\alpha \in (0, 1)$,

$$x \sim y \Rightarrow \alpha x + (1 - \alpha)z \sim \alpha y + (1 - \alpha)z.$$

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

A.4. Monotonicity For all $f, g \in \mathcal{F}$, if for all $s \in S$, $f(s) \succsim g(s)$, then $f \succsim g$.

Axiom A.1 says that the preference relation $\succsim$ should be rational. Axiom A.2 imposes von Neumann and Morgenstern's independence requirement on constant acts—the acts involving no state ambiguity. Axiom A.3 requires $\succsim$ to be continuous with respect to mixture coefficients. Axiom A.4 says that an act should be preferred to another act if it yields a better outcome in each state, where the ranking of outcomes is assumed to be state independent and induced by the preference relation over constant acts.

Besides the basic axioms, we are interested in studying preference relations that exhibit ambiguity aversion. There are two prominent definitions of ambiguity aversion in the literature. The first is due to Schmeidler (1989), which formulates ambiguity aversion as convexity of a preference relation. This is based on the understanding that a mixture of two acts that a decision maker feels indifferent about makes the outcome in each state less extreme and thus could be viewed as a hedge against ambiguity. So for an ambiguity averse decision maker, the mixed act, should be preferred to either of the two indifferent acts. We call this definition S-ambiguity aversion.

A.5.1. S-ambiguity aversion For all $f, g \in \mathcal{F}$ and all $\alpha \in (0, 1)$, if $f \sim g$, then $\alpha f + (1 - \alpha)g \succsim f$.

An alternative definition of ambiguity aversion is provided by Ghirardato and Marinacci (2002). The definition is based on a notion of comparative ambiguity aversion that they propose: A preference relation $\succsim_1$ is said to be **more ambiguity averse** than another preference relation $\succsim_2$ if for each $f \in \mathcal{F}$ and each $x \in X$, $x \succsim_2 f \implies x \succsim_1 f$.[1] Intuitively, $\succsim_1$ is more ambiguity averse than $\succsim_2$ if whenever the ambiguity is intolerable according to $\succsim_2$, it is also intolerable according to $\succsim_1$.

Based on the above notion, if a preference relation $\succsim$ is more ambiguity averse than a SEU preference relation, then it is said to be ambiguity averse by Ghirardato and Marinacci (2002). We call this definition GM-ambiguity aversion. Precisely, a SEU preference relation is a preference relation $\succsim$ such that there is a prior $p \in \Delta$ satisfying that for each pair $f, g \in \mathcal{F}$, $f \succsim g \iff \sum_{s \in S} p_s f(s) \succsim \sum_{s \in S} p_s g(s)$, and in this case, p is called the subjective prior. Following Ghirardato and Marinacci (2002), given $\succsim$, we call the priors in the set

$$\{p \in \Delta : \succsim' \text{ is a SEU preference relation with the subjective prior } p \text{ and } \succsim \text{ is more ambiguity averse than } \succsim'\}$$

the **benchmark priors**, and we call the set the **benchmark set**. Then GM-ambiguity aversion amounts to the non-emptiness of the benchmark set.

A.5.2. GM-ambiguity aversion The benchmark set is non-empty.

The two definitions of ambiguity aversion are not logically related in general. But under the other axioms and a strengthening of A.2, S-ambiguity aversion implies GM-ambiguity aversion (see Sect. 4).

Lastly, we require that there exist arbitrarily good and arbitrarily bad outcomes.

A.6. Unboundedness There are $x, y \in X$ such that $x \succ y$, and for each $\alpha \in (0, 1)$, there are $z, z' \in X$ satisfying that $\alpha z + (1-\alpha)y \succsim x$ and $y \succsim \alpha z' + (1-\alpha)x$.

Axiom A.6 is stronger than the usual non-degeneracy axiom which requires only that the decision maker is not indifferent about all outcomes. Axiom A.6 implies that the utility function representing the preference relation restricted to X is onto (see, e.g., Kopylov 2009; Maccheroni et al. 2006). In the literature, A.6 is assumed sometimes to simplify the analysis and sometimes to guarantee the uniqueness of a representation.[2] Throughout this paper, A.6 is imposed to simplify our presentation, and it is also indispensable for some of our results.[3]

3 Two generalized MEU decision rules

Let $\mathbb{R}$ be the set of real numbers and $\mathbb{R}_+$ the set of nonnegative real numbers.

---

[1] For preference relations satisfying A.1–A.4, the definition adopted here is equivalent to the original definition in Ghirardato and Marinacci (2002).

[2] See, e.g., Kopylov (2001), Maccheroni et al. (2006), Strzalecki (2011b), and Cerreia-Vioglio et al. (2011b).

[3] Our Theorem 3 relies on how a preference relation ranks the "limiting" acts that yield arbitrarily good or bad outcomes in all states.

3.1 Weighted maxmin rule

For each $u : X \to \mathbb{R}$ and each $f \in \mathcal{F}$, let $u(f) : S \to \mathbb{R}$ be the function obtained by composing u with f. For each $\varphi \in \mathbb{R}^S$ and each $p \in \Delta$, let $E_p\varphi$ denote the expected value of φ with respect to p.

Definition 1 A preference relation $\succsim$ admits a **weighted maxmin (WM) representation** if there exist an affine onto function $u : X \to \mathbb{R}$, a non-empty, closed, and convex set $D \subseteq \Delta$, and a function $\lambda : \mathbb{R}^S \to [0, 1]$ continuous on $\{\varphi \in \mathbb{R}^S : \min_{p\in D} E_p\varphi \neq \max_{p\in D} E_p\varphi\}$ such that for each pair $f, g \in \mathcal{F}$,

$$
\begin{aligned}
f \succsim g \iff & \lambda(u(f)) \min_{p\in D} E_p u(f) + (1 - \lambda(u(f))) \max_{p\in D} E_p u(f) \\
& \geq \lambda(u(g)) \min_{p\in D} E_p u(g) + (1 - \lambda(u(g))) \max_{p\in D} E_p u(g). \qquad (1)
\end{aligned}
$$

We denote the representation by $\langle u, D, \lambda\rangle$. We call D an admissible set of priors, or simply an **admissible set**, and λ a **weight function**. We call $\succsim$ a **WM preference relation**.

For example, Hurwicz's α-pessimism decision rule (Hurwicz 1951), also known as the α-MEU rule, admits a WM representation with the weight function being a constant equal to α. When $\alpha = 1$, it reduces to Gilboa and Schmeidler (1989)'s MEU decision rule. MEU preference relations exhibit S-ambiguity aversion (A.5.1). But in general, α-MEU preference relations do not exhibit S-ambiguity aversion, and thus, neither do WM preference relations.

One important issue is that the WM representation of a preference relation is typically not unique, due to the non-uniqueness of admissible sets. In fact, given a WM representation $\langle u, D, \lambda\rangle$ of $\succsim$, each closed and convex superset D' of D with $D' \subseteq \Delta$ is also an admissible set, since there is a weight function λ' such that $\langle u, D', \lambda'\rangle$ is also a WM representation of $\succsim$. This is because for each $f \in \mathcal{F}$, $[\min_{p\in D} E_p u(f), \max_{p\in D} E_p u(f)] \subseteq [\min_{p\in D'} E_p u(f), \max_{p\in D'} u(f)]$, and thus, there is $\lambda' : \mathbb{R}^S \to [0, 1]$ such that for each $f \in \mathcal{F}$,

$$
\begin{aligned}
& \lambda(u(f)) \min_{p\in D} E_p u(f) + (1 - \lambda(u(f))) \max_{p\in D} E_p u(f) \\
& \quad = \lambda'(u(f)) \min_{p\in D'} E_p u(f) + (1 - \lambda'(u(f))) \max_{p\in D'} E_p u(f).
\end{aligned}
$$

Example 1 shows that a SEU preference relation admits for each $\epsilon \in [0, 1]$, a WM representation in which the admissible set is the set of ϵ-contaminations of the subjective prior.

Example 1 [4]Assume $X := \mathbb{R}$ for simplicity. Let $p^* \in \Delta$. Consider the following SEU preference relation $\succsim$ over $\mathbb{R}^S$ with the subjective prior p^*: For each pair $f, g \in \mathbb{R}^S$, $f \succsim g \iff E_{p^*} f \geq E_{p^*} g$.

[4] I thank the referee for suggesting this example.

For each $\epsilon \in [0,1]$, let $D^\epsilon := \{(1-\epsilon)p^* + \epsilon p : p \in \Delta\}$ denote the set of ϵ-contaminations of the subjective prior p^*. Note that $D^0 = \{p^*\}$, $D^1 = \Delta$, and for each pair $\epsilon, \epsilon' \in [0,1]$ with $\epsilon \leq \epsilon'$, $D^\epsilon \subseteq D^{\epsilon'}$. Fix $\epsilon \in [0,1]$. Define $\lambda^\epsilon : \mathbb{R}^S \to \mathbb{R}$ by setting for each $f \in \mathbb{R}^S$,

$$\lambda^\epsilon(f) := \begin{cases} \dfrac{\max_{s\in S} f(s) - E_{p^*} f}{\max_{s\in S} f(s) - \min_{s\in S} f(s)} & \text{if } \min_{s\in S} f(s) \neq \max_{s\in S} f(s), \\ 1 & \text{if } \min_{s\in S} f(s) = \max_{s\in S} f(s). \end{cases}$$

Observe that for each $f \in \mathbb{R}^S$, λ^ϵ is continuous at f whenever $\min_{p\in D^\epsilon} E_p f \neq \max_{p\in D^\epsilon} E_p f$. Moreover, for each $f \in \mathbb{R}^S$,

$$E_{p^*} f = \lambda^\epsilon(f) \min_{p\in D^\epsilon} E_p f + (1 - \lambda^\epsilon(f)) \max_{p\in D^\epsilon} E_p f,$$

since $\min_{p\in D^\epsilon} E_p f = (1-\epsilon)E_{p^*} f + \epsilon \min_{s\in S} f(s)$ and $\max_{p\in D^\epsilon} E_p f = (1-\epsilon)E_{p^*} f + \epsilon \max_{s\in S} f(s)$. Let u be the identity mapping on $\mathbb{R}$. Then $\langle u, D^\epsilon, \lambda^\epsilon\rangle$ is a WM representation of $\succsim$.

In light of the nesting property of admissible sets, we are interested in finding the **smallest** one if it exists. It provides a natural normalization of WM representations. In particular, given a WM representation $\langle u, D, \lambda\rangle$ of $\succsim$, the maximum and minimum expected utilities of each act over D provide, respectively, an upper and lower bounds of the value of the act, and the bounds are tight if D is the smallest admissible set.

Definition 2 A **canonical WM representation** of $\succsim$ is a WM representation $\langle u, D, \lambda\rangle$ of $\succsim$ such that D is the smallest admissible set, i.e., for each WM representation $\langle u', D', \lambda'\rangle$ of $\succsim$, $D \subseteq D'$.

Ghirardato et al. (2004) introduce a WM representation different from a canonical WM representation. Instead of seeking for the smallest admissible set, they require the admissible set in their representation to satisfy some property related to an "unambiguous" preference relation induced from the initial preference relation. More precisely, their admissible set is required to be one component of another representation à la Bewley (2002) of the unambiguous preference relation. We therefore call their representation of the initial preference relation a **Bewley WM representation**. To define it formally, we recall Ghirardato et al. (2004)'s notion of unambiguous preference relation (see also Nehring (2007)). Given $\succsim$, for each pair $f, g \in \mathcal{F}$, f is said to be **unambiguously preferred to** g, denoted by $f \succsim^* g$, if for each $\alpha \in (0,1]$ and each $h \in \mathcal{F}$,

$$\alpha f + (1-\alpha)h \succsim \alpha g + (1-\alpha)h.$$

Intuitively, if the consideration of either hedging against or speculating on ambiguity does not affect the ranking of acts f and g, then it is as if that the decision maker unambiguously prefers f to g (see Ghirardato et al. 2004 for a discussion).[5] Moreover, recall that $\succsim^*$ admits a **Bewley representation** if there exist an affine onto function $u : X \to \mathbb{R}$ and a non-empty, closed, and convex set $D \subseteq \Delta$ such that for each pair $f, g \in \mathcal{F}$,

$$f \succsim^* g \iff \text{for each } p \in D,\ E_p u(f) \geq E_p u(g). \tag{2}$$

We denote the Bewley representation by $\langle u, D \rangle$. It can be shown that u is unique up to a positive affine transformation and D is unique.[6]

Definition 3 A **Bewley WM representation** of $\succsim$ is a WM representation $\langle u, D, \lambda \rangle$ of $\succsim$ such that $\langle u, D \rangle$ is a Bewley representation of the unambiguous preference relation $\succsim^*$ induced from $\succsim$. We call D the Bewley set of priors, or simply the **Bewley set**.

It is known that a preference relation admits a Bewley WM representation as long as it satisfies some basic axioms.

Proposition 1 *A preference relation $\succsim$ satisfies A.1–A.4 and A.6 if and only if it admits a Bewley weighted maxmin representation $\langle u, D, \lambda \rangle$. Moreover, u is unique up to a positive affine transformation, D is unique, and given u, λ is unique on $\{\varphi \in \mathbb{R}^S : \min_{p \in D} E_p \varphi \neq \max_{p \in D} E_p \varphi\}$.*[7]

A version of the "only if" direction of Proposition 1 is shown by Cerreia-Vioglio et al. (2011a). Ghirardato et al. (2004) also show a version of the "only if" direction but with the additional certainty independence axiom of Gilboa and Schmeidler (1989), so that their weight function has more structure.[8] Given the existing results, it is not hard to check the "if" direction.

In general, when a Bewley WM representation exists, a canonical WM representation may not, and even when both WM representations exist, they could be different (see Sect. 3.1.1). Our first main result says that if $\succsim$ satisfies S-ambiguity aversion in addition to A.1–A.4 and A.6, then besides admitting a Bewley WM representation, it also admits a canonical WM representation. Moreover, the two types of WM representations coincide.

The converse holds if some additional conditions on the representation are imposed. For each non-empty, closed, and convex set $D \subseteq \Delta$, let, for each $\varphi \in \mathbb{R}^S$, $l(\varphi; D) := \max_{p \in D} E_p \varphi - \min_{p \in D} E_p \varphi$, and let $\Lambda(D)$ be the set of functions $\lambda : \mathbb{R}^S \to [0, 1]$ satisfying that

(i) (monotonicity) for each pair $\varphi, \varphi' \in \mathbb{R}^S$ with $\varphi' \geq \varphi$,

$$\max_{p \in D} E_p \varphi' - \max_{p \in D} E_p \varphi \geq \lambda(\varphi') l(\varphi'; D) - \lambda(\varphi) l(\varphi; D); \tag{3}$$

---

[5] Equivalently, the unambiguous preference relation $\succsim^*$ is the maximal restriction of $\succsim$ that satisfies the independence axiom (Nehring 2007).

[6] For example, see Proposition 5 of Ghirardato et al. (2004).

[7] When $\min_{p \in D} E_p \varphi = \max_{p \in D} E_p \varphi$, the choice of $\lambda(\varphi)$ does not matter.

[8] Their weight function is constant additive and positively homogeneous of degree 1.

(ii) (quasi-concavity) for each pair $\varphi, \varphi' \in \mathbb{R}^S$ satisfying (3) and for $\varphi'' := \frac{\varphi+\varphi'}{2}$,

$$\max_{p\in D} E_p\varphi'' - \max_{p\in D} E_p\varphi \geq \lambda(\varphi'')l(\varphi''; D) - \lambda(\varphi)l(\varphi; D). \tag{4}$$

Theorem 1 *A preference relation $\succsim$ satisfies A.1–A.4, A.5.1, and A.6 if and only if it admits a canonical weighted maxmin representation $\langle u, D, \lambda\rangle$ with $\lambda \in \Lambda(D)$. Moreover, u is unique up to a positive affine transformation, D is unique and coincides with the Bewley set, and given u, λ is unique on $\{\varphi \in \mathbb{R}^S : \min_{p\in D} E_p\varphi \neq \max_{p\in D} E_p\varphi\}$.*

A well-known representation of the class of preference relations satisfying A.1–A.4, A.5.1, and A.6, different from ours, is provided by Cerreia-Vioglio et al. (2011b). To define it formally, let $\mathcal{G}$ be the set of functions $G : \mathbb{R}\times\Delta \to (-\infty, \infty]$ such that (i) G is quasi-convex and lower semicontinuous on $\mathbb{R}\times\Delta$, (ii) for each $p \in \Delta$, $G(\cdot, p)$ is non-decreasing on $\mathbb{R}$, and (iii) for each $t \in \mathbb{R}$, $\min_{p\in\Delta} G(t, p) = t$. A function $G \in \mathcal{G}$ is said to be linearly continuous if the functional $I : \mathbb{R}^S \to \mathbb{R}$, defined by setting for each $\varphi \in \mathbb{R}^S$, $I(\varphi) := \min_{p\in\Delta} G(E_p\varphi, p)$, is continuous.

Definition 4 A preference relation $\succsim$ admits an **ambiguity averse representation** if there exist an affine onto function $u : X \to \mathbb{R}$ and a linearly continuous function $G \in \mathcal{G}$ such that for each pair $f, g \in \mathcal{F}$,

$$f \succsim g \iff \min_{p\in\Delta} G(E_p u(f), p) \geq \min_{p\in\Delta} G(E_p u(g), p)$$

We denote the representation by $\langle u, G\rangle$.

It is shown by Cerreia-Vioglio et al. (2011b) that a preference relation satisfies A.1–A.4, A.5.1, and A.6 if and only if it admits an ambiguity averse representation $\langle u, G\rangle$, where u is unique up to a positive affine transformation, and given u, G is unique.[9] Given a preference relation $\succsim$ that admits an ambiguity averse representation $\langle u, G\rangle$, they define the set[10]

$$D^* := cl(\{p \in \Delta : \text{ for some } t \in \mathbb{R}, G(t, p) < \infty\}), \tag{5}$$

and show that $\langle u, D^*\rangle$ is a Bewley representation of the unambiguous preference relation $\succsim^*$ induced from $\succsim$.[11]

Our result above, in view of theirs, not only provides an alternative representation of such a preference relation, but also shows that the set D^* derived from their representation is the smallest admissible set.

[9] See Theorems 3 and 5 and Proposition 4 of Cerreia-Vioglio et al. (2011b).

[10] The set D^* is independent of the choice of ambiguity averse representations. Indeed, by Proposition 4 of Cerreia-Vioglio et al. (2011b), if $\langle u', G'\rangle$ is another ambiguity averse representation of $\succsim$, then for each $p \in \Delta$, $G(t, p) < \infty$ for some $t \in \mathbb{R}$ if and only if $G'(t', p) < \infty$ for some $t' \in \mathbb{R}$.

[11] See Proposition 9 and Theorem 10 of Cerreia-Vioglio et al. (2011b).

Corollary 1 *If a preference relation admits an ambiguity averse representation* $\langle u, G\rangle$*, then it also admits a canonical weighted maxmin representation* $\langle u, D^*, \lambda\rangle$ *with* $\lambda \in \Lambda(D^*)$.

3.1.1 Relation between a canonical WM representation and a Bewley WM representation

Restricted to preference relations satisfying S-ambiguity aversion, we have shown that a canonical WM representation is the same as a Bewley WM representation. But this is not the case for preference relations that do not satisfy S-ambiguity aversion. We now provide two examples, respectively, showing that (1) a preference relation admitting a Bewley WM representation may not admit a canonical WM representation, and (2) even if a preference relation admits both types of representations, the Bewley set may not be the smallest admissible set. The proofs of our claims in the two examples are given in "Appendix."

Example 2 **(The Bewley set exists, but the smallest admissible set does not)** Let $S := \{1, 2, 3\}$ and $X := \mathbb{R}$ for simplicity. Let $p' := (\frac{2}{3}, \frac{1}{12}, \frac{1}{4})$, $q' := (\frac{1}{12}, \frac{2}{3}, \frac{1}{4})$, and $p^* := (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ be probabilities over S, where the sth coordinate denotes the probability of state s, $s \in S$. Let $D_1 := \{p', q'\}$ and $D_2 := \{p \in \Delta : d(p, p^*) \leq \frac{1}{\sqrt{6}}\}$. Define $V : \mathbb{R}^S \to \mathbb{R}$ by setting for each $f \in \mathbb{R}^S$,

$$V(f) := \begin{cases} \min\limits_{p\in D_1} E_p f & \text{if } \max\{f(1), f(2)\} < f(3), \\ \min\limits_{p\in D_2} E_p f & \text{if } \max\{f(1), f(2)\} \geq f(3). \end{cases}$$

Define $\succsim$ over $\mathbb{R}^S$ by setting for each pair $f, g \in \mathbb{R}^S$, $f \succsim g \iff V(f) \geq V(g)$.

The preference relation $\succsim$ satisfies A.1–A.4 and A.6. By Proposition 1, $\succsim$ admits a Bewley WM representation.[12] But $\succsim$ does not admit a canonical WM representation.

Example 3 **(The Bewley set is not the smallest admissible set)** Let S, X, p', q', p^*, and D_2 be defined as in Example 2. Let $p'' := (\frac{2}{3}, \frac{1}{6}, \frac{1}{6})$ and $q'' := (\frac{1}{6}, \frac{2}{3}, \frac{1}{6})$ be probabilities over S. For each $f \in \mathbb{R}^S$ with $\max\{f(1), f(2)\} < f(3)$, let

$$\alpha(f) := \begin{cases} \text{med}\left\{0, \dfrac{1 - E_{p''}f}{1 - E_{p''}f + E_{p'}f}, 1\right\} & \text{if } f(1) \leq f(2) < f(3), \\ \text{med}\left\{0, \dfrac{1 - E_{q''}f}{1 - E_{q''}f + E_{q'}f}, 1\right\} & \text{if } f(2) < f(1) < f(3), \end{cases}$$

and

$$D_1(f) := \{\alpha(f)p' + (1 - \alpha(f))p'', \alpha(f)q' + (1 - \alpha(f))q''\},$$

12 It can be shown that the Bewley set is $cl(co(D_1 \cup \{p \in D_2 | p_3 \geq \frac{1}{6}\}))$.

where med is the median operator.[13] Define $V : \mathbb{R}^S \to \mathbb{R}$ by setting for each $f \in \mathbb{R}^S$,

$$V(f) := \begin{cases} \min_{p \in D_1(f)} E_p f & \text{if } \max\{f(1), f(2)\} < f(3), \\ \min_{p \in D_2} E_p f & \text{if } \max\{f(1), f(2)\} \geq f(3). \end{cases}$$

Define $\succsim$ over $\mathbb{R}^S$ by setting for each pair $f, g \in \mathbb{R}^S$, $f \succsim g \iff V(f) \geq V(g)$.

Like Example 2, $\succsim$ satisfies A.1–A.4 and A.6 and thus admits a Bewley WM representation.[14] Moreover, $\succsim$ admits a canonical WM representation in which the smallest admissible set is $\{p \in D_2 : p_3 \geq \frac{1}{6}\}$. But this set is not the Bewley set.

3.2 Variant constraint rule

For each pair $p, q \in \Delta$, let $d(p, q)$ denote the Euclidean distance between p and q. For each $p \in D$ and each closed subset of A of Δ, define the distance between p and A, denoted by $d(p, A)$, to be $\min_{q \in A} d(p, q)$. Given $\succsim$, for each $f \in \mathcal{F}$, let $x_{*f} \in X$ denote the worst outcome yielded by f, i.e., $x_{*f} = f(s)$ for some $s \in S$ such that for each $s' \in S$, $f(s') \succsim x_{*f}$.

Definition 5 A preference relation $\succsim$ admits a **variant constraint (VC) representation** if there exist an affine onto function $u : X \to \mathbb{R}$, a non-empty, closed, and convex set $K \subseteq \Delta$, and a function $\sigma : \mathbb{R}^S \to \mathbb{R}_+$ continuous on $\{u(f) \in \mathbb{R}^S : f \in \mathcal{F}, f \nsim x_{*f}\}$ and lower semicontinuous on $\{u(f) \in \mathbb{R}^S : f \in \mathcal{F}, f \sim x_{*f}\}$ such that for each pair $f, g \in \mathcal{F}$,

$$f \succsim g \iff \min_{p \in \Delta : d(p,K) \leq \sigma(u(f))} E_p u(f) \geq \min_{p \in \Delta : d(p,K) \leq \sigma(u(g))} E_p u(g).$$

We denote the representation by $\langle u, K, \sigma \rangle$. We call K an essential set of priors, or simply an **essential set**, and σ a **constraint function**. We call $\succsim$ a **VC preference relation**.

A MEU preference relation admits a VC representation in which σ is a constant equal to 0. While MEU preference relations exhibit S-ambiguity aversion, this is not true for VC preference relations in general. Instead, VC preference relations satisfy GM-ambiguity aversion. Moreover, we will show that VC preference relations are characterized by GM-ambiguity aversion together with some basic axioms.

Like a WM representation, a VC representation of a preference relation is typically not unique, due to the non-uniqueness of essential sets. In fact, given a VC representation $\langle u, K, \sigma \rangle$ of $\succsim$, each non-empty, closed, and convex subset K' of K is an essential set, since there is a constraint function σ' such that $\langle u, K', \sigma' \rangle$ is also a VC representation of $\succsim$. Example 4 illustrates this fact using a MEU preference relation.

---

13 For each $t \in \mathbb{R}$, med$\{0, t, 1\}$ is the median of $0, t, 1$.

14 It can be shown that the Bewley set is the same as in Example 2.

Example 4 Let S, X, p^*, D_2 be defined as in Example 2. Consider the MEU preference relation $\succsim$ over $\mathbb{R}^S$ defined by setting for each pair $f, g \in \mathbb{R}^S$, $f \succsim g \iff \min_{p \in D_2} E_p f \geq \min_{p \in D_2} E_p g$.

For each $\epsilon \in [0, \frac{1}{\sqrt{6}}]$, let $K^\epsilon := \{p \in \Delta : d(p, p^*) \leq \epsilon\}$. Note that $K^0 = \{p^*\}$, $K^{\frac{1}{\sqrt{6}}} = D_2$, and for each pair $\epsilon, \epsilon' \in [0, \frac{1}{\sqrt{6}}]$ with $\epsilon \leq \epsilon'$, $K^\epsilon \subseteq K^{\epsilon'}$. Fix $\epsilon \in [0, \frac{1}{\sqrt{6}}]$. Define $\sigma^\epsilon : \mathbb{R}^S \to \mathbb{R}_+$ by setting for each $f \in \mathbb{R}^S$, $\sigma^\epsilon(f) := \frac{1}{\sqrt{6}} - \epsilon$. Observe that σ^ϵ is continuous, and that for each $f \in \mathbb{R}^S$,

$$\min_{p \in D_2} E_p f = \min_{p \in \Delta : d(p, K^\epsilon) \leq \sigma^\epsilon(f)} E_p f,$$

since $D_2 = \{p \in \Delta : d(p, K^\epsilon) \leq \sigma^\epsilon(f)\}$. Let u be the identity mapping on $\mathbb{R}$. Then $\langle u, K^\epsilon, \sigma^\epsilon \rangle$ is a VC representation of $\succsim$.

In light of the nesting property of essential sets, we can define a canonical VC representation in a similar way as defining a canonical WM representation.

Definition 6 A **canonical VC representation** of $\succsim$ is a VC representation $\langle u, K, \sigma \rangle$ of $\succsim$ such that K is the largest essential set, i.e., for each VC representation $\langle u', K', \sigma' \rangle$ of $\succsim$, $K' \subseteq K$.

Our second main result says that if $\succsim$ satisfies GM-ambiguity aversion in addition to A.1–A.4 and A.6, then it admits a canonical VC representation. Moreover, the largest essential set is the set of benchmark priors.

The converse holds with an additional condition imposed on the VC representation. For each non-empty, closed, and convex set $K \subseteq \Delta$, let $\Sigma(K)$ denote the set of functions $\sigma : \mathbb{R}^S \to \mathbb{R}_+$ such that for each pair $\varphi, \varphi' \in \mathbb{R}^S$ with $\varphi' \geq \varphi$,

$$\min_{p \in \Delta : d(p, K) \leq \sigma(\varphi')} E_p \varphi' \geq \min_{p \in \Delta : d(p, K) \leq \sigma(\varphi)} E_p \varphi. \qquad (6)$$

Theorem 2 *A preference relation satisfies A.1–A.4, A.5.2, and A.6 if and only if it admits a canonical variant constraint representation $\langle u, K, \sigma \rangle$ with $\sigma \in \Sigma(K)$. Moreover, u is unique up to a positive affine transformation, K is unique and coincides with the benchmark set, and given u, σ is unique on $\{u(f) \in \mathbb{R}^S : f \in \mathcal{F}, f \nsim x_{*f}\}$.*[15]

While Cerreia-Vioglio et al. (2011b) provide the characterization of preference relations exhibiting S-ambiguity aversion, Theorem 2 provides the characterization of preference relations exhibiting GM-ambiguity aversion.

The class of preference relations admitting canonical VC representations is different from the class admitting canonical WM representations. In view of Theorems 1 and 2, there are preference relations admitting canonical WM representations but not canonical VC representations, since S-ambiguity aversion does not imply GM-ambiguity

[15] If $x_{*f} \sim f$, then $\min_{s \in S} u(f(s)) = u(x_f)$. Thus as long as $\sigma(u(f))$ is sufficiently large, $\min_{p \in D : d(p,K) \leq \sigma(u(f))} E_p u(f) = u(x_f)$.

aversion. On the other hand, there are preference relations admitting canonical VC representations but not canonical WM representations, as shown in Example 5.

Example 5 Let $\succsim$ be defined as in Example 2. As discussed in Example 2, $\succsim$ satisfies A.1–A.4 and A.6, and $\succsim$ does not admit a canonical WM representation. Since for each $f \in \mathbb{R}^S$, $E_{p^*} f \geq V(f)$, it can be readily seen that p^* is a benchmark prior. Then $\succsim$ satisfies A.5.2, and thus, by Theorem 2, $\succsim$ admits a canonical VC representation.[16]

When $\succsim$ admits both a canonical WM representation and a canonical VC representation, the largest essential set is always a subset of the smallest admissible set, and the two sets coincide if and only if $\succsim$ is a MEU preference relation.

Proposition 2 *If a preference relation admits both a canonical weighted maxmin representation $\langle u, D, \lambda\rangle$ and a canonical variant constraint representation $\langle u', K, \sigma\rangle$, then $K \subseteq D$.*

Proposition 3 *A preference relation is a MEU preference relation if and only if it admits both a canonical weighted maxmin representation $\langle u, D, \lambda\rangle$ and a canonical variant constraint representation $\langle u', K, \sigma\rangle$ with $K = D$, λ being a constant function equal to 1, and σ being a constant function equal to 0.*

Assume that $\succsim$ satisfies A.1–A.4, A.5.1, and A.6. Recall that by Cerreia-Vioglio et al. (2011b), $\succsim$ admits an ambiguity averse representation $\langle u, G\rangle$. We have shown in Sect. 3.1 that $\succsim$ also admits a canonical WM representation in which D^* defined by (5) is the smallest admissible set. A similar result can be obtained here. Formally, let

$$K^* := \{p \in \Delta : \text{ for each } t \in \mathbb{R}, G(t, p) = t\}. \tag{7}$$

Like D^*, the set K^* is independent of the choice of ambiguity averse representations of $\succsim$. It can be readily shown that K^* is the benchmark set. Thus, $\succsim$ satisfies GM-ambiguity aversion if and only if K^* is non-empty. When K^* is non-empty, by Theorem 2, $\succsim$ admits a canonical VC representation in which K^* is the largest essential set.

Corollary 2 *If a preference relation admits an ambiguity averse representation $\langle u, G\rangle$ and $K^* \neq \emptyset$, then it also admits a canonical variant constraint representation $\langle u, K^*, \sigma\rangle$ with $\sigma \in \Sigma(K^*)$.*

3.3 Different representations of the multiplier decision rule

There is a large class of preference relations exhibiting both S-ambiguity aversion and GM-ambiguity aversion as well as satisfying A.1–A.4 and A.6. Thus, besides admitting an ambiguity aversion representation studied by Cerreia-Vioglio et al. (2011b), this class of preference relations also admits a canonical WM representation and a canonical VC representation. Although the three representations have different behavioral interpretations, they are observationally equivalent.

[16] It can be shown that the largest essential set is $\{p \in D_2 : p_3 \geq \frac{1}{4}\}$.

On the other hand, different representations of a decision rule facilitate our understanding of the rule. As an example, we consider two important robust decision rules introduced by Hansen and Sargent (2001): the constraint rule and the multiplier rule.[17] While the multiplier rule is originally defined in terms of an ambiguity averse representation, we will provide two alternative representations. We will argue that each of them provides a different way of understanding the connection between the two decision rules.

For each pair $p, q \in \Delta$, we write $p \ll q$ if p is absolutely continuous with respect to q, and we denote the relative entropy of p with respect to q by $R(p||q)$, i.e.,

$$R(p||q) = \begin{cases} \sum_{s \in S} p_s \log \dfrac{p_s}{q_s} & \text{if } p \ll q, \\ \infty & \text{otherwise.} \end{cases}$$

The relative entropy, known as Kullback–Leibler divergence, is a measure of "distance" between two probabilities.

Definition 7 A preference relation $\succsim$ admits a **constraint representation** if there exist an affine onto function $u : X \to \mathbb{R}$, a prior $q \in \Delta$, and a constant $\eta \in \mathbb{R}_+$ such that for each pair $f, g \in \mathcal{F}$,

$$f \succsim g \iff \min_{p \in \Delta : R(p||q) \le \eta} E_p u(f) \ge \min_{p \in \Delta : R(p||q) \le \eta} E_p u(g). \tag{8}$$

We denote the representation by $\langle u, q, \eta \rangle$. We call $\succsim$ a **constraint preference relation**.

A constraint representation $\langle u, q, \eta \rangle$ is a special WM representation $\langle u, D, \lambda \rangle$ in which $D = \{p \in \Delta : R(p||q) \le \eta\}$ and λ is a constant function equal to 1. In spirit, it can also be viewed as a special VC representation with a singleton essential set, an alternative measure of "distance" between two probabilities, and a constant constraint function.

Definition 8 A preference relation $\succsim$ admits a **multiplier representation** if there exist an affine onto function $u : X \to \mathbb{R}$, a prior $q \in \Delta$, and a constant $\theta \in (0, \infty]$ such that for each pair $f, g \in \mathcal{F}$,

$$f \succsim g \iff \min_{p \in \Delta} [E_p u(f) + \theta R(p||q)] \ge \min_{p \in \Delta} [E_p u(g) + \theta R(p||q)]. \tag{9}$$

We denote the representation by $\langle u, q, \theta \rangle$. We call $\succsim$ a **multiplier preference relation**.[18]

---

[17] Strzalecki (2011a) axiomatizes the multiplier rule.

[18] The preference relation in Definition 8 is called a multiplier preference relation since the parameter θ in the unconstrained minimization problem in (9) can be viewed as a Lagrange multiplier in the Lagrangian of the constrained minimization problem in (8).

A multiplier representation $\langle u, q, \theta \rangle$ is an ambiguity averse representation $\langle u, G \rangle$ in which $G : \mathbb{R} \times \Delta \to (-\infty, \infty]$ is given by, for each $(t, p) \in \mathbb{R} \times \Delta$, $G(t, p) = t + \theta R(p||q)$. A multiplier preference relation exhibits both S-ambiguity aversion and GM-ambiguity aversion and satisfies A.1–A.4 and A.6. By Theorems 1 and 2, it admits both a canonical WM representation and a canonical VC representation. Moreover, we can explicitly pin down the components of each representation.

Proposition 4 *Assume that a preference relation admits a multiplier representation $\langle u, q, \theta \rangle$. Then it admits a canonical weighted maxmin representation $\langle u, D, \lambda \rangle$. When $\theta < \infty$, $D = \{p \in \Delta : p \ll q\}$, and for each $\varphi \in \mathbb{R}^S$,*

$$\lambda(\varphi) = \begin{cases} \dfrac{\max\limits_{s \in S: q_s > 0} \varphi(s) + \theta \log E_q e^{-\frac{\varphi}{\theta}}}{\max\limits_{s \in S: q_s > 0} \varphi(s) - \min\limits_{s \in S: q_s > 0} \varphi(s)} & \text{if } \min\limits_{s \in S: q_s > 0} \varphi(s) < \max\limits_{s \in S: q_s > 0} \varphi(s), \\ 1 & \text{if } \min\limits_{s \in S: q_s > 0} \varphi(s) = \max\limits_{s \in S: q_s > 0} \varphi(s), \end{cases} \tag{10}$$

and when $\theta = \infty$, $D = \{q\}$, and λ is a constant function equal to 1. It also admits a canonical variant constraint representation $\langle u, K, \sigma \rangle$. When $\theta < \infty$, $K = \{q\}$, and for each $\varphi \in \mathbb{R}^S$,

$$\sigma(\varphi) = \min_{p \in \Delta : E_p \varphi = -\theta \log E_q e^{-\frac{\varphi}{\theta}}} d(p, q), \tag{11}$$

and when $\theta = \infty$, $K = \{q\}$, and σ is a constant function equal to 0.

A connection between the multiplier rule and the constraint rule is established by Hansen and Sargent (2001, 2008) in a dynamic resource allocation problem. Fixing $u : X \to \mathbb{R}$ and $q \in \Delta$, they show, under some conditions, that for each η, there is θ such that the constraint rule $\langle u, q, \eta \rangle$ and the multiplier rule $\langle u, q, \theta \rangle$ generate the same optimal allocation, and vise versa. But in general, as they point out, the two decision rules induce totally different preference relations.

A further understanding of their connection can be obtained in view of each of our alternative representations. First, both constraint and multiplier preference relations are WM preference relations, and the smallest admissible sets and weight functions are different. The smallest admissible set of a constraint preference relation is a neighborhood of a central prior with the "distance" measure being the relative entropy and the radius a constant. In contrast, the smallest admissible set of a multiplier preference relation is either the set of all priors absolutely continuous with respect to a central prior or the set containing only the central prior. More importantly, a constraint preference relation always evaluates an act by its worst possible expected utility. In contrast, a multiplier preference relation evaluates an act by a weighted average of its best and worst possible expected utilities, and the weights vary across acts.

Second, both constraint and multiplier preference relations are VC preference relations, and the largest essential sets and constraint functions are different. While the largest essential set of a constraint preference relation is a neighborhood of a central prior, that of a multiplier preference relation is the set containing only a central prior. More importantly, although both preference relations evaluate an act by its worst possible expected utility, the set of possible priors stays constant for a constraint preference relation while it changes across acts for a multiplier preference relation. This can be seen more transparently when we replace the Euclidean distance measure with relative entropy in the VC representation of the multiplier rule. Indeed, if $\succsim$ admits a multiplier representation $\langle u, q, \theta \rangle$ with $\theta < \infty$, then for each pair $f, g \in \mathcal{F}$,

$$f \succsim g \iff \min_{p \in \Delta : R(p||q) \leq \sigma'(u(f))} E_p u(f) \geq \min_{p \in \Delta : R(p||q) \leq \sigma'(u(g))} E_p u(g) \tag{12}$$

where $\sigma' : \mathbb{R}^S \to \mathbb{R}$ is given by, for each $\varphi \in \mathbb{R}^S$,

$$\sigma'(\varphi) = \min_{p \in \Delta : E_p \varphi = -\theta \log E_q e^{-\frac{\varphi}{\theta}}} R(p||q).$$

Comparing (8) and (12), we can see that a constraint preference relation always regards the priors in a fixed neighborhood of q possible, whereas a multiplier preference relation considers neighborhoods of different sizes for different acts.

Note that the key implication of using relative entropy instead of the Euclidean distance in a VC representation is that it excludes the priors that are not absolutely continuous with respect to a central prior. Indeed, only when $p \ll q$, $R(p||q) < \infty$. But this is not an issue for a multiplier preference relation, since by definition (9), it disregards all the priors that are not absolutely continuous with respect to the central prior q. Thus, relative entropy can be used instead of the Euclidean distance in the VC representation of a multiplier preference, as long as the constraint function is adjusted correspondingly.

4 Changing ambiguity aversion

Within the class of preference relations exhibiting S-ambiguity aversion and GM-ambiguity aversion and satisfying A.1–A.4 and A.6, our goal, in this section, is to study those exhibiting particular patterns of changing ambiguity aversion with respect to changes in wealth. We provide a definition of decreasing/increasing absolute ambiguity aversion and investigate its implication on the representations studied in Sect. 3.

4.1 Wealth effect and decreasing absolute ambiguity aversion

To motivate the study of the wealth effect, consider the following variation of Ellsberg (1961)'s thought experiment.

Example 6 An urn contains 100 balls, of which 33 are red, and 67 are either black or white. A ball is drawn from the urn. For each $t \in \mathbb{R}_+$, r_t denotes the act "betting on red." It pays $100 + t$ dollars if the ball is red and t dollars otherwise. Let b_t denote the act "betting on black," and its payoff is similarly given. See the payoff Tables 1, 2 and 3.

Table 1 Payoffs of r_t and b_t

$t \in \mathbb{R}_+$	Red	Black	White
r_t	$100 + t$	t	t
b_t	t	$100 + t$	t

Table 2 Payoffs of r_0 and b_0

$t = 0$	Red	Black	White
r_0	100	0	0
b_0	0	100	0

Table 3 Payoffs of r_{10^4} and b_{10^4}

$t = 10^4$	Red	Black	White
r_{10^4}	10,100	10,000	10,000
b_{10^4}	10,000	10,100	10,000

Assume that a decision maker's preference relation $\succsim$ satisfies Axioms A.1–A.4, A.5.1, and A.6, and assume for simplicity that he is risk neutral. For each $t \in \mathbb{R}_+$, r_t is an unambiguous act which yields $100 + t$ with probability 0.33 and t with probability 0.67, whereas b_t is an ambiguous act which yields $100 + t$ and t with unknown probabilities. The decision maker may prefer r_0 to b_0 if he is averse to ambiguity, while the degree of his ambiguity aversion may decrease with the increase in the baseline prize t. It can be expected that when t is sufficiently large, he is willing to take the ambiguity bearing act, and prefers say b_{10^4} to r_{10^4}.

Such a behavioral pattern is evidenced by the laboratory experiments of Baillon and Placido (2015) with subjects that may or may not be risk neutral. In the following, we propose a set of behavioral axioms to capture this and other analogous phenomena. The experiments designed by Baillon and Placido (2015) basically test the implications of our axioms.

A.2.1. Decreasing absolute ambiguity aversion.[19] For all $f \in \mathcal{F}$, $x, y, z \in X$ and $\alpha \in (0, 1)$, if either f is a constant act or $y \succsim x$, then

$$\begin{aligned} &\alpha f + (1-\alpha)x \succsim \alpha z + (1-\alpha)x \\ &\quad \Rightarrow \alpha f + (1-\alpha)y \succsim \alpha z + (1-\alpha)y. \end{aligned} \tag{13}$$

When f is constant, (13) is essentially von Neumann–Morgenstern's independence requirement on constant acts. When f is not constant, then (13) says that if an (ambiguous) act $\alpha f + (1-\alpha)x$ is preferred to a constant act $\alpha z + (1-\alpha)x$, then it is still the case after improving the certainty part from x to y for both acts. In other words, if the

[19] After completing this paper, I learned that Ghirardato and Siniscalchi independently propose a very similar axiom of decreasing absolute ambiguity aversion in their work "Symmetric preferences," presented in RUD 2015 and D-TEA 2015.

ambiguity is previously tolerable, it is even more tolerable after a common improvement on the certainty part. Axiom A.2.1 implies that if an act induces a larger utility than another act by an ensured amount t in each state, then the value of the former act is larger than the later by at least t.

Similarly, if we replace $y \succsim x$ in A.2.1 by $x \succsim y$, then $\succsim$ exhibits increasing absolute ambiguity aversion.

A.2.2. Increasing absolute ambiguity aversion For all $f \in \mathcal{F}$, $x, y, z \in X$ and $\alpha \in (0, 1)$, if either f is a constant act or $x \succsim y$, then

$$\alpha f + (1-\alpha)x \succsim \alpha z + (1-\alpha)x \\ \Rightarrow \alpha f + (1-\alpha)y \succsim \alpha z + (1-\alpha)y.$$

If we require both A.2.1 and A.2.2 to hold, then $\succsim$ exhibits constant absolute ambiguity aversion defined by Grant and Polak (2013).

A.2.3. Constant absolute ambiguity aversion (Grant and Polak 2013) For all $f \in \mathcal{F}$, $x, y, z \in X$ and $\alpha \in (0, 1)$,

$$\alpha f + (1-\alpha)x \succsim \alpha z + (1-\alpha)x \\ \Rightarrow \alpha f + (1-\alpha)y \succsim \alpha z + (1-\alpha)y. \tag{14}$$

While the above axioms deal with the effect of an absolute change in the certainty part of an act, one can imagine a similar effect of a relative change in the proportion of the certainty part of an act. We refer the readers to Maccheroni et al. (2006) for such a thought experiment. To capture the increase in ambiguity aversion in the relative size of ambiguity, we propose the following axiom as analogous to A.2.2.

A.2.4. Increasing relative ambiguity aversion For all $f \in \mathcal{F}$, $x, z \in X$ and $\alpha, \beta \in (0, 1)$, if $\alpha \geq \beta$, then

$$\alpha f + (1-\alpha)x \succsim \alpha z + (1-\alpha)x \\ \Rightarrow \beta f + (1-\beta)x \succsim \beta z + (1-\beta)x. \tag{15}$$

Axiom A.2.4 says that if $\alpha f + (1-\alpha)x$ is preferred to a constant act $\alpha z + (1-\alpha)x$, then this is still the case after the proportion of the certainty part increases in both acts. In other words, the degree of ambiguity aversion increases as the proportion of the ambiguous part of an act increases. Similarly, a preference relation exhibits **decreasing relative ambiguity aversion** if (15) holds for each pair $\alpha, \beta \in (0, 1)$ such that $\alpha \leq \beta$, and **constant relative ambiguity aversion** if (15) holds for each pair $\alpha, \beta \in (0, 1)$. Chateauneuf and Faro (2009) propose a so-called worst independence axiom under the assumption that there is a worst outcome in X. Their axiom amounts to our constant relative ambiguity aversion for the class of preference relations satisfying A.1, A.3, A.4, A.5.1, and exhibiting constant absolute ambiguity aversion. In general, for preference relations satisfying A.1–A.4, their axiom is implied by ours.

It is known that under A.1 and A.3–A.6, preference relations exhibiting constant absolute ambiguity aversion also exhibit increasing relative ambiguity aversion. We will show that decreasing absolute ambiguity aversion also implies increasing relative ambiguity aversion in some limit form.

4.2 Characterizations

For each $t \in \mathbb{R}$, we denote by $t\mathbf{1}$ a constant element in $\mathbb{R}^S$ with each coordinate being t, and when $t = 1$, we simply write $\mathbf{1}$. For each non-empty, closed, and convex set $K \subseteq \Delta$, let $\bar{\Sigma}(K)$ be the set of functions $\sigma \in \Sigma(K)$ such that for each pair $\varphi, \varphi' \in \mathbb{R}^S$ satisfying (6) and for $\varphi'' := \frac{\varphi+\varphi'}{2}$,

$$\min_{p \in \Delta : d(p,K) \leq \sigma(\varphi'')} E_p \varphi'' \geq \min_{p \in \Delta : d(p,K) \leq \sigma(\varphi)} E_p \varphi.$$

That is, the constraint functions in $\bar{\Sigma}(K)$ satisfy not only the monotonicity requirement imposed on the functions in $\Sigma(K)$, but also a quasi-concavity requirement.

Theorem 3 *Let a preference relation $\succsim$ be given. The following statements are equivalent.*

1. *The preference relation $\succsim$ satisfies A.1, A.2.1, A.3, A.4, A.5.1, and A.6.*
2. *The preference relation $\succsim$ admits a weighted maxmin representation $\langle u, D, \lambda \rangle$ such that $\lambda \in \Lambda(D)$, and for each $\varphi \in \mathbb{R}^S$, $\lambda(\varphi + t\mathbf{1})$ is non-increasing in t and* $\lim_{k\to\infty} \lim_{t\to\infty} \lambda(k\varphi - t\mathbf{1}) = 1$.
3. *The preference relation $\succsim$ admits a variant constraint representation $\langle u, K, \sigma \rangle$ such that $\sigma \in \bar{\Sigma}(K)$, and for each $\varphi \in \mathbb{R}^S$, $\sigma(\varphi + t\mathbf{1})$ is non-increasing in t and* $\lim_{k\searrow 0} \lim_{t\to\infty} \sigma(k\varphi + t\mathbf{1}) = 0$.
4. *The preference relation $\succsim$ admits an ambiguity averse representation $\langle u, G \rangle$ such that for each $p \in \Delta$, $G(t, p) - t$ is non-decreasing in t.*

Moreover, $\langle u, D, \lambda \rangle$ is a canonical weighted maxmin representation, $\langle u, K, \sigma \rangle$ is a canonical variant constraint representation, and for each $\varphi \in \mathbb{R}^S$, $\lim_{t\to\infty} \lambda(k\varphi - t\mathbf{1})$ *and* $\lim_{t\to\infty} \sigma(k\varphi + t\mathbf{1})$ *is non-decreasing in k on* $(0, \infty)$.

Decreasing absolute ambiguity aversion has straightforward behavioral implications on the WM representation and the VC representation of a preference relation. As the baseline utility of an act increases, it is as if that a WM decision maker puts less weight on the worst case, and a VC decision maker considers a smaller neighborhood of approximating priors.

Note that under A.1, A.2.1, A.3, and A.4, S-ambiguity aversion (A.5.1) implies GM-ambiguity aversion (A.5.2). Thus, by Theorems 1 and 2, the preference relation admits both a canonical WM representation and a canonical VC representation. However, the equivalence of statements 1, 2, and 3 is **not** a corollary of Theorems 1 and 2. In Theorem 3, we do not require the admissible set in the WM representation to be the smallest. Neither do we require the essential set in the VC representation to be the largest. Instead, we impose a limit condition on each of the representations, which turns out to determine the representation uniquely. In fact, the two limit conditions are characterizing conditions of the smallest admissible set and the largest essential set, respectively. Hence, both representations are canonical.

The two limit conditions have natural interpretations. The condition $\lim_{k\to\infty}\lim_{t\to\infty}\lambda(k\varphi - t\mathbf{1}) = 1$ says that a WM decision maker tends to consider only the worst case in an extremely bad situation where the baseline utility of an act is sufficiently low and the scale of its ambiguous part is sufficiently large. The condition $\lim_{k\searrow 0}\lim_{t\to\infty}\sigma(k\varphi + t\mathbf{1}) = 0$ says that a VC decision maker tends to consider only approximating priors in the essential set K in an extremely good situation where the baseline utility of an act is sufficiently high and the scale of its ambiguous part is sufficiently small.

Theorem 3 also shows that decreasing absolute ambiguity aversion implies increasing relative ambiguity aversion in some limit form under the other axioms. The fact that $\lim_{t\to\infty}\lambda(k\varphi - t\mathbf{1})$ is non-decreasing in k on $(0, \infty)$ implies that when the baseline utility of an act is sufficiently low, if the scale of its ambiguous part increases, then a WM decision maker behaves as if he is more pessimistic and puts a larger weight on the worst case. Similarly, $\lim_{t\to\infty}\sigma(k\varphi + t\mathbf{1})$ being non-decreasing in k on $(0, \infty)$ implies that when the baseline utility of an act is sufficiently high, if the scale of its ambiguous part increases, then a VC decision maker behaves as if he is more cautious and considers a larger neighborhood of approximating priors.

Analogous representations can be obtained for preference relations exhibiting increasing absolute ambiguity aversion: Theorem 3 holds if A.2.1 is replaced by A.2.2 and t by $-t$. Since A.2.3 is equivalent to the combination of A.2.1 and A.2.2, we further obtain the representations for preference relations exhibiting constant absolute ambiguity aversion.

Corollary 3 *Let a preference relation $\succsim$ be given. The following statements are equivalent.*

1. *The preference relation $\succsim$ satisfies A.1, A.2.3, A.3, A.4, A.5.1, and A.6.*
2. *The preference relation $\succsim$ admits a weighted maxmin representation $\langle u, D, \lambda\rangle$ such that $\lambda \in \Lambda(D)$, and for each $\varphi \in \mathbb{R}^S$, $\lambda(\varphi + t\mathbf{1})$ is constant in t and $\lim_{k\to\infty}\lambda(k\varphi) = 1$.*
3. *The preference relation $\succsim$ admits a variant constraint representation $\langle u, K, \sigma\rangle$ such that $\sigma \in \bar{\Sigma}(K)$, and for each $\varphi \in \mathbb{R}^S$, $\sigma(\varphi + t\mathbf{1})$ is constant in t and $\lim_{k\searrow 0}\sigma(k\varphi) = 0$.*
4. *The preference relation $\succsim$ admits an ambiguity averse representation $\langle u, G\rangle$ such that for each $p \in \Delta$, $G(t, p) - t$ is constant in t.*

Moreover, $\langle u, D, \lambda\rangle$ is a canonical weighted maxmin representation; $\langle u, K, \sigma\rangle$ is a canonical variant constraint representation, and for each $\varphi \in \mathbb{R}^S$, $\lambda(k\varphi)$ and $\sigma(k\varphi)$ is non-decreasing in k on $(0, \infty)$.

As shown in Corollary 3, constant absolute ambiguity aversion implies that no matter how the baseline utility of an act changes, a WM decision maker puts the same weight on the worst case, and a VC decision maker considers the same neighborhood of approximating priors.

It is known that constant absolute ambiguity aversion implies increasing relative ambiguity aversion under the other axioms in Corollary 3. We show that when the scale of an (ambiguous) act increases, it is as if that a WM decision maker puts a larger weight on the worst case, or that a VC decision maker considers a larger neighborhood of approximating priors.

The equivalence of statements 1 and 4 in Corollary 3 reproduces some existing results in the literature. Maccheroni et al. (2006) propose a weak certainty independence axiom and call a preference relation a **variational preference relation** if it satisfies A.1, A.3–A.6, and weak certainty independence. They show that a preference relation is a variational preference relation if and only if there exist an affine onto function $u : X \to \mathbb{R}$ and a lower semicontinuous convex function $c : \Delta \to [0, \infty]$ with $\min_{p \in \Delta} c(p) = 0$ such that

$$f \succsim g \iff \min_{p \in \Delta}[E_p u(f) + c(p)] \geq \min_{p \in \Delta}[E_p u(g) + c(p)]. \tag{16}$$

Cerreia-Vioglio et al. (2011b) show that this representation amounts to an ambiguity averse representation $\langle u, G \rangle$ in which G is additively separable.[20] Since weak certainty independence is known to be equivalent to constant absolute ambiguity aversion under the other axioms,[21] our result can be expected since G is additively separable if for each $p \in \Delta$, $G(t, p) - t$ is constant in t.

We close this subsection by providing a differential characterization of the smallest admissible set for preference relations exhibiting decreasing absolute ambiguity aversion. Let a preference relation $\succsim$ be given. Following Rigotti and Strazalecki (2008), define the correspondence $\pi : \mathcal{F} \rightrightarrows \Delta$ by setting for each $f \in \mathcal{F}$,

$$\pi(f) := \left\{ p \in \Delta : \sum_{s \in S} p_s f(s) \succsim \sum_{s \in S} p_s g(s) \Longrightarrow f \succsim g \right\}. \tag{17}$$

The interpretation is that the set $\pi(f)$ consists of all the prior beliefs that rationalize the choice of f over other acts (see Cerreia-Vioglio et al. 2011b). These beliefs are useful in the study of ambiguity averse preference relations and their applications (see, e.g., Rigotti and Strazalecki 2008; Lang 2017). Mathematically, they correspond to the supporting hyperplanes of the upper contour set of f.

For a general preference relation satisfying A.1–A.4, A.5.1, and A.6, Cerreia-Vioglio et al. (2011b) show that $D^* = cl(co(\bigcup_{f \in \mathcal{F}} \pi(f)))$. Thus, by our Corollary 1, $cl(co(\bigcup_{f \in \mathcal{F}} \pi(f)))$ is the smallest admissible set. Proposition 5 strengthens this result for preference relations exhibiting decreasing absolute ambiguity aversion.

Proposition 5 *Assume that a preference relation $\succsim$ satisfies A.1, A.2.1, A.3, A.4, A.5.1, and A.6. Let $\langle u, D, \lambda \rangle$ be a canonical weighted maxmin representation of $\succsim$. Then for each $x \in X$, $D = cl(co(\bigcup_{f \sim x} \pi(f)))$.*

Proposition 5 implies that to identify the smallest admissible set, it suffices to find the set of beliefs that rationalize the decision maker's choices of the acts that lie on one indifference curve.

---

[20] The function $G : \mathbb{R} \times \Delta \to (-\infty, \infty]$ is additively separable if there are $\gamma : \mathbb{R} \to \mathbb{R}$ and $c : \Delta \to [0, \infty]$ such that for each $(t, p) \in \mathbb{R} \times \Delta$, $G(t, p) = \gamma(t) + c(p)$. See their Proposition 12.

[21] Grant and Polak (2013) show the equivalence under A.1, A.3, and weaker versions of A.4 and A.6.

Note that a differential characterization of the largest essential set can be obtained by restricting π to constant acts. For each $x \in X$, the set $\pi(x)$ is interpreted as the set of beliefs that rationalize the choice of the constant act x over ambiguous acts. It is known that the benchmark set coincides with $\bigcap_{x \in X} \pi(x)$ (see, e.g., Cerreia-Vioglio et al. 2011a). Thus, by Theorem 2, for each preference relation that admits a canonical VC representation, the largest essential set is $\bigcap_{x \in X} \pi(x)$.

Lang (2017) assumes that X is a set of simple lotteries over $\mathbb{R}$ and proposes definitions, for an ambiguity averse preference relation, to distinguish between first-order and second-order ambiguity aversion at a wealth level $w \in \mathbb{R}$. Lang (2017) characterizes second-order ambiguity aversion at w by $\pi(w)$ being a singleton. In view of Lang (2017)'s result, for an ambiguity averse preference relation that exhibits second-order ambiguity aversion at each $w \in \mathbb{R}$, it admits a canonical VC representation only if there is $p \in \Delta$ such that for each $w \in \mathbb{R}$, $\pi(w) = \{p\}$. In this case, the largest essential set is exactly the singleton $\{p\}$.

4.3 Comparison with other definitions

Based on a notion of comparative "dispersion," Chambers et al. (2014) also propose the definitions of decreasing, increasing, and constant absolute ambiguity aversion. According to them, an act f is considered **at least as dispersed as** an act g, denoted by $f \trianglerighteq g$, if there exist $x \in X$ and $\lambda \in [0, 1]$ such that $g = \lambda f + (1 - \lambda)x$. They say that a preference relation $\succsim$ exhibits decreasing (increasing) absolute ambiguity aversion if for each pair $f, g \in \mathcal{F}$ with $f \trianglerighteq g$, each pair $x, y \in X$ with $y \succsim x$ ($x \succsim y$), and each $\alpha \in (0, 1)$,

$$\begin{aligned} &\alpha f + (1-\alpha)x \succsim \alpha g + (1-\alpha)x \\ &\quad \Rightarrow \alpha f + (1-\alpha)y \succsim \alpha g + (1-\alpha)y. \end{aligned} \tag{18}$$

They say that $\succsim$ exhibits constant absolute ambiguity aversion if it exhibits both their decreasing and increasing absolute ambiguity aversion.

Since for each $f \in \mathcal{F}$ and each $z \in X$, $f \trianglerighteq z$, their decreasing absolute ambiguity aversion is stronger than ours. If we apply their definition, then we get a similar result as Theorem 3 with modified monotonicity conditions of λ, σ and G.[22] This is also the case for increasing absolute ambiguity aversion. As for constant absolute ambiguity aversion, it turns out that their seemingly stronger definition is in fact equivalent to A.2.3.

Proposition 6 *Assume that a preference relation $\succsim$ satisfies A.1–A.4 and A.6. Then for each pair $f, g \in \mathcal{F}$ with $f \trianglerighteq g$, each pair $x, y \in X$, and each $\alpha \in (0, 1)$, (18) holds if and only if $\succsim$ satisfies A.2.3.*

---

[22] If $\succsim$ satisfies A.1–A.4 and A.6, then there exist an affine and onto function $u : X \to \mathbb{R}$, and a non-decreasing and continuous functional $I : \mathbb{R}^S \to \mathbb{R}$ such that $f \succsim g \iff I(u(f)) \geq I(u(g))$. Their decreasing absolute ambiguity aversion amounts to that for each pair $\varphi, \varphi' \in \mathbb{R}^S$, if $\varphi' = \lambda\varphi + (1-\lambda)t'$ for some $\lambda \in [0, 1]$ and some $t' \in \mathbb{R}$, and if $I(\varphi) = I(\varphi')$, then for each $t > 0$, $I(\varphi + t\mathbf{1}) \geq I(\varphi' + t\mathbf{1})$, which means, in the WM representation for example, that $\lambda(\varphi + t\mathbf{1}) \leq \lambda(\varphi' + t\mathbf{1})$.

The validity of our axioms can be seen from their implications in the smooth ambiguity model of Klibanoff et al. (2005). A preference relation $\succsim$ admits a **smooth ambiguity representation** if there exist an affine onto function $u : X \to \mathbb{R}$, an increasing function $\phi : \mathbb{R} \to \mathbb{R}$, a countably additive Borel probability measure μ over Δ such that for each pair $f, g \in \mathcal{F}$,

$$f \succsim g \Longleftrightarrow \int_{p\in\Delta} \phi(E_p u(f)) d\mu(p) \geq \int_{p\in\Delta} \phi(E_p u(g)) d\mu(p). \tag{19}$$

We denote such a representation by $\langle u, \phi, \mu \rangle$.

Klibanoff et al. (2005) show that in this model "attitudes toward pure risk are characterized by the shape of u, as usual, while attitudes toward ambiguity are characterized by the shape of ϕ," and "one advantage of this model is that the well-developed machinery for dealing with risk attitudes can be applied as well to ambiguity attitudes." We focus on the case where ϕ is concave so that $\langle u, \phi, \mu \rangle$ represents a preference relation exhibiting S-ambiguity aversion. In this case, our definitions of decreasing and increasing absolute ambiguity aversion correspond exactly to the usual monotonicity properties of the Arrow–Pratt coefficient of absolute risk aversion of ϕ (Arrow 1965; Pratt 1964).

Proposition 7 *Fix an affine onto function $u : \mathbb{R} \to \mathbb{R}$ and an increasing, concave, and twice differentiable function $\phi : \mathbb{R} \to \mathbb{R}$. Each preference relation represented by $\langle u, \phi, \mu \rangle$ for some countably additive Borel probability measure μ on Δ satisfies A.2.1 (A.2.2) if and only if ϕ exhibits decreasing (increasing) absolute risk aversion, i.e., $-\frac{\phi''}{\phi'}$ is non-increasing (non-decreasing).*

As a corollary of Proposition 7, we obtain Cerreia-Vioglio et al. (2011b)'s characterization of constant absolute ambiguity aversion in terms of constant absolute risk aversion of ϕ (see their Theorem 23).

Cherbonniera and Gollier (2015) propose a definition of decreasing aversion under ambiguity in the smooth ambiguity model. They assume that the decision maker with initial wealth $z \in \mathbb{R}_+$ is facing N possible monetary lotteries $(\tilde{x}_1, \ldots, \tilde{x}_N)$. For each $n \in \{1, \ldots, N\}$, $\tilde{x}_n$ occurs with probability q_n. The value function of the decision maker obeys the smooth ambiguity rule:

$$\sum_{n=1}^{N} q_n \phi(Eu(z + \tilde{x}_n)),$$

where ϕ is increasing and concave. According to them, the decision maker exhibits decreasing aversion if

$$\phi^{-1}\left(\sum_{n=1}^{N} q_n \phi(Eu(z + \tilde{x}_n))\right) = u(z)$$
$$\Rightarrow \sum_{n=1}^{n} q_n \phi'(Eu(z + \tilde{x}_n)) Eu'(z + \tilde{x}_n) \geq \phi'(u(z))u'(z). \tag{20}$$

The key difference between their definition and ours is that their definition does not distinguish the effect of wealth on risk aversion and ambiguity aversion, while ours captures the effect of baseline utilities on ambiguity aversion. More precisely, their definition says that an ambiguous lottery becomes more desirable at a higher **monetary wealth** level, while our axiom essentially says that it becomes more desirable at a higher **baseline utility** level. When the comparison of behavior is based on changes in baseline utilities, we are restricted to the effect of baseline utilities on ambiguity aversion.

Indeed, they show that (20) holds if and only if both u and $\phi \circ u$ exhibit decreasing concavity, where u summarizes the decision maker's risk attitude according to Klibanoff et al. (2005). Instead, as shown in Proposition 7, our A.2.1 corresponds only to the decreasing concavity of ϕ, the measure of the decision maker's ambiguity attitude.

Cherbonniera and Gollier (2015) also provide an analogous definition for decreasing aversion in the α-maxmin expected utility model studied by Ghirardato et al. (2004). That is,

$$\alpha \min_n Eu(z+\tilde{x}_n) + (1-\alpha)\max_n Eu(z+\tilde{x}_n) \le u(z)$$
$$\Rightarrow \forall z' \le z,\ \alpha \min_n Eu(z'+\tilde{x}_n) + (1-\alpha)\max_n Eu(z'+\tilde{x}_n) \le u(z'). \qquad (21)$$

With the weight α being fixed, their definition only imposes restriction on the function u which summarizes the decision maker's risk attitude according to Ghirardato et al. (2004). In contrast, our axiom captures the decision maker's changing ambiguity attitudes, which is reflected by assigning less weight on the worst case as the baseline utility of an act increases (see Theorem 3).

Focusing also on the effect of changes in monetary wealth, Cerreia-Vioglio et al. (2017) provide a definition of decreasing/increasing absolute ambiguity aversion in a general setting in which X is assumed to be a set of monetary lotteries. Given a lottery x and a wealth level w, the transformed lottery at w, x^w is defined as a lottery that yields a payoff of $c+w$ with the same probability as x yields c. Intuitively, x^w is the "real" lottery faced by a decision maker at the wealth level w. Given a preference relation $\succsim$ and a wealth level w, they define the induced preference relation $\succsim^w$ at w as a preference relation that ranks acts as the initial preference relation $\succsim$ ranks "real" acts that yield in each state transformed lotteries at w. Then based on the notion of comparative ambiguity aversion introduced by Ghirardato and Marinacci (2002), a preference relation $\succsim$ is said to exhibit decreasing absolute ambiguity aversion if for each pair w, w' with $w' > w$, $\succsim^w$ is more ambiguity averse than $\succsim^{w'}$.

One implication of their definition is that if a preference relation exhibits decreasing absolute ambiguity aversion, then it must exhibit constant absolute risk aversion. Their definition does not allow, for example, a decision maker to exhibit both decreasing absolute ambiguity aversion and decreasing absolute risk aversion. In contrast, our definition does not impose restrictions on a decision maker's risk attitude and captures changing ambiguity aversion with respect to changes in baseline utilities. Thus, we can accommodate the possibility that a decision maker exhibits both decreasing absolute ambiguity aversion and decreasing absolute risk aversion. In case a decision maker is risk neutral, their definition has the same implication as ours on the representations, since changes in monetary wealth translate directly to changes in baseline utilities.

5 Conclusion

In this paper, we study two extensions of the well-known MEU decision rule to accommodate a decision maker's changing ambiguity attitude: a WM rule and a VC rule. Due to the non-uniqueness of their representations, we are interested in finding canonical representations of the two rules in terms of the smallest admissible set and the largest essential set, respectively. We characterize a class of preference relations that admits a canonical WM representation as well as a class that admits a canonical VC representation. The first class of preference relations exhibits S-ambiguity aversion, while the second exhibits GM-ambiguity aversion. In the second part of this paper, we study the wealth effect under ambiguity. We propose axioms of decreasing (increasing and constant) absolute and relative ambiguity aversion. Representations are provided for the subclass of ambiguity averse preference relations exhibiting decreasing (increasing and constant) absolute ambiguity aversion. The monotonic pattern of changing ambiguity aversion is reflected in a intuitive way in both the WM representation and the VC representation.

Appendix: Proofs

We denote by $\mathbb{N}$ the set of positive integers. For each $\varphi \in \mathbb{R}^S$, let $\varphi^* := \max_{s \in S} \varphi(s)$ and $\varphi_* := \min_{s \in S} \varphi(s)$. A functional $I : \mathbb{R}^S \to \mathbb{R}$ is **normalized** if for each $t \in \mathbb{R}$, $I(t\mathbf{1}) = t$. It is **constant additive** if for each $\varphi \in \mathbb{R}^S$ and each $t \in \mathbb{R}$, $I(\varphi + t\mathbf{1}) = I(\varphi) + t$. It is **constant superadditive** if for each $\varphi \in \mathbb{R}^S$ and each $t \in \mathbb{R}_+$, $I(\varphi + t\mathbf{1}) \geq I(\varphi) + t$. Lastly, it is **superadditive** if for each pair $\varphi, \varphi' \in \mathbb{R}^S$, $I(\varphi + \varphi') \geq I(\varphi) + I(\varphi')$.

Proofs in Section 3

Lemma 1 *A preference relation $\succsim$ satisfies A.1–A.4 and A.6 if and only if there exist an affine onto function $u : X \to \mathbb{R}$ and a normalized, non-decreasing, and continuous functional $I : \mathbb{R}^S \to \mathbb{R}$ such that for each pair $f, g \in \mathcal{F}$,*

$$f \succsim g \iff I(u(f)) \geq I(u(g)). \tag{22}$$

Moreover, u is unique up to a positive affine transformation, and given u, there is a unique normalized functional $I : \mathbb{R}^S \to \mathbb{R}$ satisfying (22).

Proof To prove the sufficiency of the axioms, let $\succsim$ satisfy A.1–A.4 and A.6. Note that A.6 implies the usual non-degeneracy axiom: There are $f, g \in \mathcal{F}$ such that

$f \succ g$. Then by applying the same techniques of Lemma 57 of Cerreia-Vioglio et al. (2011b), one can show that there exist a non-constant affine function $u : X \to \mathbb{R}$ and a normalized, non-decreasing, and continuous functional $I : u(X)^S \to \mathbb{R}$ satisfying (22). Although A.5.1 is assumed in their Lemma 57, the axiom is only used to show that I is quasi-concave, and dropping it does not affect the existence of u and I satisfying the other properties. Moreover, since $\succsim$ satisfies A.6, by Lemma 29 of Maccheroni et al. (2006), $u(X) = \mathbb{R}$.

The necessity of the axioms can be readily seen, and the uniqueness follows from routine arguments. Thus, we omit the proofs. □

Proof of Proposition 1 To prove the sufficiency of the axioms, let $\succsim$ satisfy A.1–A.4 and A.6. By Lemma 1, there exist an affine onto function $u : X \to \mathbb{R}$ and a normalized, non-decreasing, and continuous functional $I : \mathbb{R}^S \to \mathbb{R}$ satisfying (22). Then by Propositions 1 and 5 of Cerreia-Vioglio et al. (2011a) and their proof of Proposition 5,[23] there exist a non-empty, closed, and convex set $D \subseteq \Delta$ and a function $\lambda : \mathbb{R}^S \to \mathbb{R}$ such that (i) $\langle u, D\rangle$ is a Bewley representation of the unambiguous preference relation $\succsim^*$ induced from $\succsim$, and (ii) for each $\varphi \in \mathbb{R}^S$,

$$I(\varphi) = \lambda(\varphi) \min_{p\in D} E_p\varphi + (1 - \lambda(\varphi)) \max_{p\in D} E_p\varphi. \tag{23}$$

In particular, for each $\varphi \in \mathbb{R}^S$ such that $\min_{p\in D} E_p\varphi \neq \max_{p\in D} E_p\varphi$,

$$\lambda(\varphi) = \frac{\max_{p\in D} E_p\varphi - I(\varphi)}{\max_{p\in D} E_p\varphi - \min_{p\in D} E_p\varphi}. \tag{24}$$

Since I is continuous, λ is continuous on $\{\varphi \in \mathbb{R}^S : \min_{p\in D} E_p\varphi \neq \max_{p\in D} E_p\varphi\}$. Thus, it can be readily seen that $\langle u, D, \lambda\rangle$ is a Bewley WM representation of $\succsim$.

To prove the necessity of the axioms, let $\succsim$ admit a Bewley WM representation $\langle u, D, \lambda\rangle$. Clearly, $\succsim$ satisfies A.1. Since u is affine, it satisfies A.2. Since u is onto, by Lemma 29 of Maccheroni et al. (2006), it satisfies A.6.

To show that $\succsim$ satisfies A.3, define $J : \mathbb{R}^S \to \mathbb{R}$ by setting for each $\varphi \in \mathbb{R}^S$,

$$J(\varphi) := \lambda(\varphi) \min_{p\in D} E_p\varphi + (1 - \lambda(\varphi)) \max_{p\in D} E_p\varphi,$$

and it suffices to show that J is continuous. Let $\varphi \in \mathbb{R}^S$ and $\{\varphi^n\}_{n=1}^\infty$ be a sequence of elements in $\mathbb{R}^S$ such that $\lim_{n\to\infty} \varphi^n = \varphi$. If $\min_{p\in D} E_p\varphi < \max_{p\in D} E_p\varphi$, then λ is continuous at φ, and thus, $\lim_{n\to\infty} J(\varphi^n) = J(\varphi)$. If $\min_{p\in D} E_p\varphi = \max_{p\in D} E_p\varphi$, then $\lim_{n\to\infty} \min_{p\in D} E_p\varphi^n = \lim_{n\to\infty} \max_{p\in D} E_p\varphi^n = J(\varphi)$, and thus, $\lim_{n\to\infty} J(\varphi^n) = J(\varphi)$.

[23] See also Propositions 4 and 5 of Ghirardato et al. (2004).

To show that $\succsim$ satisfies A.4, let $f, g \in \mathcal{F}$ be such that for each $s \in S$, $f(s) \succsim g(s)$, and we want to show that $f \succsim g$. Since for each $s \in S$, $f(s) \succsim g(s)$, and since $\langle u, D, \lambda \rangle$ is a Bewley WM representation of $\succsim$, $u(f) \geq u(g)$. Then for each $p \in D$, $E_p u(f) \geq E_p u(g)$. Besides, by the definition of a Bewley WM representation, $\langle u, D \rangle$ is a Bewley representation of the unambiguous preference relation $\succsim^*$ induced from $\succsim$. Hence, $f \succsim^* g$. Thus, by the definition of $\succsim^*$, $f \succsim g$.

Lastly, to prove the uniqueness of the representation,[24] let $\langle u, D, \lambda \rangle$ and $\langle u', D', \lambda' \rangle$ be two Bewley WM representations of some $\succsim$. Since both u and u' are affine functions representing $\succsim$ restricted to X, by routine arguments, u' is a positive affine transformation of u. Moreover, by the definition of a Bewley WM representation, both $\langle u, D \rangle$ and $\langle u', D' \rangle$ are Bewley representations of the unambiguous preference relation $\succsim^*$ induced from $\succsim$. By Proposition 5 of Ghirardato et al. (2004), $D = D'$. Finally, suppose that $u = u'$, let $\varphi \in \mathbb{R}^S$ be such that $\min_{p \in D} E_p \varphi \neq \max_{p \in D} E_p \varphi$, and we want to show that $\lambda(\varphi) = \lambda'(\varphi)$. Let $f \in \mathcal{F}$ be such that $u(f) = \varphi$. Since $x_f \sim f$ and both $\langle u, D, \lambda \rangle$ and $\langle u', D', \lambda' \rangle$ are WM representations of $\succsim$,

$$u(x_f) = \lambda(u(f)) \min_{p \in D} E_p u(f) + (1 - \lambda(u(f))) \max_{p \in D} E_p u(f), \tag{25}$$

and

$$u'(x_f) = \lambda(u'(f)) \min_{p \in D'} E_p u'(f) + (1 - \lambda(u'(f))) \max_{p \in D'} E_p u'(f). \tag{26}$$

Since $u = u'$, $u(f) = \varphi$, and $D = D'$, by (25) and (26),

$$\lambda(\varphi) \min_{p \in D} E_p \varphi + (1 - \lambda(\varphi)) \max_{p \in D} E_p \varphi = \lambda'(\varphi) \min_{p \in D} E_p \varphi + (1 - \lambda'(\varphi)) \max_{p \in D} E_p \varphi. \tag{27}$$

Since $\max_{p \in D} E_p \varphi \neq \min_{p \in D} E_p \varphi$, by (27), $\lambda(\varphi) = \lambda'(\varphi)$. □

Proof of Theorem 1 To show the sufficiency of the axioms, let $\succsim$ satisfy A.1–A.4, A.5.1, and A.6. By the proof of Proposition 1, it admits a Bewley WM representation $\langle u, D, \lambda \rangle$, with $I : \mathbb{R}^S \to \mathbb{R}$ given by (23) being non-decreasing and continuous. We first check that $\lambda \in \Lambda(D)$. Let $\varphi, \varphi' \in \mathbb{R}^S$. If $\varphi' \geq \varphi$, then by the monotonicity of I, $I(\varphi') \geq I(\varphi)$, which implies (3). To check (4), suppose that φ, φ' satisfy (3), and let $\varphi'' := \frac{\varphi + \varphi'}{2}$. Since φ, φ' satisfy (3), $I(\varphi') \geq I(\varphi)$. Thus, by the monotonicity and continuity of I, there is $t \in \mathbb{R}_+$ such that $I(\varphi' - t\mathbf{1}) = I(\varphi)$. Let $f, g \in \mathcal{F}$ be such that $u(f) = \varphi$ and $u(g) = \varphi' - t\mathbf{1}$. Then $I(u(f)) = I(u(g))$, so $f \sim g$. Since $f \sim g$, by A.5.1, $\frac{1}{2} f + \frac{1}{2} g \succsim f$. Thus, $I(\frac{1}{2}\varphi + \frac{1}{2}(\varphi' - t\mathbf{1})) = I(u(\frac{1}{2} f + \frac{1}{2} g)) \geq I(u(f)) = I(\varphi)$. Since I is non-decreasing, $I(\frac{1}{2}\varphi + \frac{1}{2}\varphi') \geq I(\frac{1}{2}\varphi + \frac{1}{2}(\varphi' - t\mathbf{1}))$. Thus, $I(\varphi'') \geq I(\varphi)$, which implies (4).

---

[24] The uniqueness property here differs slightly from that in Cerreia-Vioglio et al. (2011a). They state the uniqueness of the representation under an additional condition that they are interested in (condition (iii) of their Proposition 5), while the uniqueness is actually guaranteed without imposing this condition.

We then prove that $\langle u, D, \lambda\rangle$ is also a canonical WM representation of $\succsim$. That is, let $\langle u', D', \lambda'\rangle$ be another WM representation of $\succsim$ and we shall prove that $D \subseteq D'$. Since $\succsim$ satisfies A.1–A.6, by Theorems 3 and 5 of Cerreia-Vioglio et al. (2011b), it admits an ambiguity averse representation $\langle v, G\rangle$, where $G : \mathbb{R} \times \Delta \to (-\infty, \infty]$ is given by, for each $(t, p) \in \mathbb{R} \times \Delta$,

$$G(t, p) = \sup\{v(x_f) : f \in \mathcal{F}, E_p v(f) \le t\}. \tag{28}$$

Recall the set D^* defined in (5). By Proposition 9 and Theorem 10 of Cerreia-Vioglio et al. (2011b), $\langle v, D^*\rangle$ is a Bewley representation of the unambiguous preference relation $\succsim^*$ induced from $\succsim$. On the other hand, by the definition of a Bewley WM representation, $\langle u, D\rangle$ is also a Bewley representation of $\succsim^*$. Thus, by Proposition 5 of Ghirardato et al. (2004), $D = D^*$. Hence, to show that $D \subseteq D'$, it is equivalent to show that $D^* \subseteq D'$.

Suppose to the contrary that $D^* \not\subseteq D'$. Since $D^* = cl(\{p \in \Delta : G(t, p) < \infty \text{ for some } t \in \mathbb{R}\})$ and D' is closed, there exist $q \in \Delta\backslash D'$ and $t \in \mathbb{R}$ such that $G(t, q) < \infty$. Then by a standard separation theorem, there is $\varphi \in \mathbb{R}^S$ such that $E_q\varphi < 0 < \min_{p\in D'} E_p\varphi$. Let $x \in X$ be such that $v(x) = t$. Since $E_q v(x) = t$, by (28), $G(t, q) \ge v(x) = t$. Let $n \in \mathbb{N}$ be such that $E_q n\varphi < t \le G(t, q) < \min_{p\in D'} E_p n\varphi$. Let $g \in \mathcal{F}$ be such that $v(g) = n\varphi$. Since $E_q v(g) = E_q n\varphi < t$, by (28), $v(x_g) \le G(t, q)$. Observe that $\min_{p\in D'} E_p u'(g) \le \lambda'(u'(g)) \min_{p\in D'} E_p u'(g) + (1-\lambda'(u'(g))) \max_{p\in D'} E_p u'(g) = u'(x_g)$. Since both v and u' are affine functions representing $\succsim$ restricted to X, by routine arguments, v is a positive affine transformation of u'. Then $\min_{p\in D'} E_p v(g) \le v(x_g)$. Recall that $G(t, q) < \min_{p\in D'} E_p n\varphi = \min_{p\in D'} E_p v(g)$. Thus, $G(t, q) < v(x_g)$, which contradicts $v(x_g) \le G(t, q)$, as desired.

To show the necessity of the axioms, let $\succsim$ admit a canonical WM representation $\langle u, D, \lambda\rangle$ with $\lambda \in \Lambda(D)$. Clearly, it satisfies A.1. Since $\lambda \in \Lambda(D)$, one can readily verify that it satisfies A.4 and A.5.1. Recall that Proposition 1 shows the necessity of A.2, A.3, and A.6 for a preference relation to admit a Bewley WM representation. Since the arguments there rely on the same properties of u, D, and λ as we have here, they can be used to show the necessity of the axioms for a preference relation to admit a canonical WM representation.

Lastly, we show the uniqueness of a canonical WM representation $\langle u, D, \lambda\rangle$ of some preference relation $\succsim$. By the definition of a canonical WM representation, D is the smallest admissible set and thus unique. As argued in the proof of the sufficiency of the axioms, D coincides with the Bewley set. Moreover, the uniqueness of u and λ follows from the same arguments as used for a Bewley WM representation in the proof of Proposition 1. □

Proof of Corollary 1 Let $\succsim$ admit an ambiguity averse representation $\langle u, G\rangle$. Define $I : \mathbb{R}^S \to \mathbb{R}$ by setting for each $\varphi \in \mathbb{R}^S$, $I(\varphi) := \min_{p\in\Delta} G(E_p\varphi, p)$. By the proof of Theorem 3 of Cerreia-Vioglio et al. (2011b), I is normalized, non-decreasing, and continuous. Moreover, by their Theorem 3, $\succsim$ satisfies A.5.1. Thus, by the proofs of

Proposition 1 and Theorem 1, $\succsim$ admits a canonical WM representation $\langle u, D^*, \lambda \rangle$ in which λ belongs to $\Lambda(D^*)$ and is given by (24) on $\{\varphi \in \mathbb{R}^S : \min_{p \in D^*} E_p \varphi \neq \max_{p \in D^*} E_p \varphi\}$. □

Proposition 8 *The preference relation $\succsim$ in Example 2 admits a Bewley weighted maxmin representation but not a canonical weighted maxmin representation.*

Proof We first prove that $\succsim$ admits a Bewley WM representation. By Proposition 1, it is equivalent to prove that $\succsim$ satisfies A.1–A.4 and A.6. Clearly, it satisfies A.1. Since for each $t \in \mathbb{R}$, $V(t\mathbf{1}) = t$, it satisfies A.2 and A.6.

To show that $\succsim$ satisfies A.3, it suffices to show the continuity of V. Let $f \in \mathbb{R}^S$ and $\{f^n\}_{n=1}^{\infty}$ be a sequence of elements in $\mathbb{R}^S$ that converges to f. We shall show that $\lim_{n\to\infty} V(f^n) = V(f)$ in each of the following two cases.

Case 1: $\max\{f(1), f(2)\} \neq f(3)$. $\max\{f(1), f(2)\} < f(3)$. Then $V(f) = \min_{p \in D_1} E_p f$. Moreover, for sufficiently large n, $\max\{f^n(1), f^n(2)\} < f^n(3)$, so that $V(f^n) = \min_{p \in D_1} E_p f^n$. Hence, $\lim_{n\to\infty} V(f^n) = \lim_{n\to\infty} \min_{p \in D_1} E_p f^n = \min_{p \in D_1} E_p f = V(f)$. Similarly, one can show that when $\max\{f(1), f(2)\} > f(3)$, $\lim_{n\to\infty} V(f^n) = V(f)$.

Case 2: $\max\{f(1), f(2)\} = f(3)$. Then $V(f) = \min_{p \in D_2} E_p f$. Suppose that $f(1) < f(2)$. Thus, $f(1) < f(2) = f(3)$ and for sufficiently large n, $f^n(1) < f^n(2)$. Recall $p'' = (\frac{2}{3}, \frac{1}{6}, \frac{1}{6})$. Then $p'' \in D_2$. For each $p \in D_2$, since

$$\frac{1}{6} \geq \left(p_1 - \frac{1}{3}\right)^2 + \left(p_2 - \frac{1}{3}\right)^2 + \left(p_3 - \frac{1}{3}\right)^2 \geq \left(p_1 - \frac{1}{3}\right)^2 + 2\left(\frac{1-p_1}{2} - \frac{1}{3}\right)^2$$
$$= \frac{3}{2}\left(p_1 - \frac{1}{3}\right)^2, \tag{29}$$

$p_1 \leq \frac{2}{3} = p''_1$. Since $f(1) < f(2) = f(3)$, and since $p'' \in D_2$ and for each $p \in D_2$, $p_1 \leq p''_1$, $\min_{p \in D_2} E_p f = E_{p''} f = E_{p'} f$. Then $V(f) = \min_{p \in D_2} E_p f = E_{p'} f$, and thus,

$$\lim_{n\to\infty} E_{p'} f^n = \lim_{n\to\infty} \min_{p \in D_2} E_p f^n = V(f). \tag{30}$$

For sufficiently large n, since $f^n(1) < f^n(2)$, and since $p'_1 > q'_1$ and $p'_3 = q'_3$, $\min_{p \in D_1} E_p f^n = E_{p'} f^n$, and thus,

$$\text{either } V(f^n) = \min_{p \in D_1} E_p f^n = E_{p'} f^n \text{ or } V(f^n) = \min_{p \in D_2} E_p f^n. \tag{31}$$

By (30) and (31), $\lim_{n\to\infty} V(f^n) = V(f)$. Similarly, one can show that when $f(1) > f(2)$, $\lim_{n\to\infty} V(f^n) = V(f)$. Suppose that $f(1) = f(2)$. Then $f(1) = f(2) = f(3)$, and thus, $\lim_{n\to\infty} \min_{p \in D_1} E_p f^n = \lim_{n\to\infty} \min_{p \in D_2} E_p f^n = \min_{p \in D_2} E_p f =$

$V(f)$. For each $n \in \mathbb{N}$, either $V(f^n) = \min_{p \in D_1} E_p f^n$ or $V(f^n) = \min_{p \in D_2} E_p f^n$. Thus, $\lim_{n\to\infty} V(f^n) = V(f)$.

Lastly, to show that $\succsim$ satisfies A.4, let $f, g \in \mathbb{R}^S$ be such that $f \geq g$. We shall show that $V(f) \geq V(g)$ in each of the following three cases.

Case 1: Either $\max\{f(1), f(2)\} < f(3)$ and $\max\{g(1), g(2)\} < g(3)$, or $\max\{f(1), f(2)\} \geq f(3)$ and $\max\{g(1), g(2)\} \geq g(3)$. Then either $V(f) = \min_{p \in D_1} E_p f$ and $V(g) = \min_{p \in D_1} E_p g$, or $V(f) = \min_{p \in D_2} E_p f$ and $V(g) = \min_{p \in D_2} E_p g$. In either scenario, since $f \geq g$, $V(f) \geq V(g)$.

Case 2: $\max\{f(1), f(2)\} < f(3)$ and $\max\{g(1), g(2)\} \geq g(3)$. Then $V(f) = \min_{p \in D_1} E_p f$ and $V(g) = \min_{p \in D_2} E_p g$. Let $f' \in \mathbb{R}^S$ be such that

$$f'(1) = f(1), \quad f'(2) = f(2), \quad f'(3) = \max\{f(1), f(2)\}.$$

Since $\max\{f(1), f(2)\} < f(3)$, $f'(3) < f(3)$. Thus, $f' \leq f$. Since $f \geq g$, $f'(3) = \max\{f(1), f(2)\} \geq \max\{g(1), g(2)\} \geq g(3)$. Thus, $f' \geq g$. Consider the sequence $\{\frac{1}{n} f + \frac{n-1}{n} f'\}_{n=1}^{\infty}$ of elements of $\mathbb{R}^S$. For each $n \in \mathbb{N}$, since $\max\{\frac{1}{n} f(1) + \frac{n-1}{n} f'(1), \frac{1}{n} f(2) + \frac{n-1}{n} f'(2)\} = \max\{f(1), f(2)\} < \frac{1}{n} f(3) + \frac{n-1}{n} f'(3)$, $V(\frac{1}{n} f + \frac{n-1}{n} f') = \min_{p \in D_1} E_p(\frac{1}{n} f + \frac{n-1}{n} f')$. Since V is continuous (as shown when proving that A.3 holds), $V(f') = \lim_{n\to\infty} V(\frac{1}{n} f + \frac{n-1}{n} f') = \lim_{n\to\infty} \min_{p \in D_1} E_p(\frac{1}{n} f + \frac{n-1}{n} f') = \min_{p \in D_1} E_p f'$. Since $f' \leq f$, $\min_{p \in D_1} E_p f' \leq \min_{p \in D_1} E_p f$. Thus, $V(f') \leq V(f)$. Since $\max\{f'(1), f'(2)\} = f'(3)$ and $f' \geq g$, $V(f') = \min_{p \in D_2} E_p f' \geq \min_{p \in D_2} E_p g = V(g)$. Hence, $V(f) \geq V(f') \geq V(g)$.

Case 3: $\max\{f(1), f(2)\} \geq f(3)$ and $\max\{g(1), g(2)\} < g(3)$. Then $V(f) = \min_{p \in D_2} E_p f$ and $V(g) = \min_{p \in D_1} E_p g$. Let $f' \in \mathbb{R}^S$ be such that

$$f'(1) = \begin{cases} f(3) & \text{if } f(1) \geq f(2), \\ f(1) & \text{if } f(1) < f(2), \end{cases} \quad f'(2) = \begin{cases} f(2) & \text{if } f(1) \geq f(2), \\ f(3) & \text{if } f(1) < f(2), \end{cases} \quad f'(3) = f(3).$$

Since $\max\{f(1), f(2)\} \geq f(3)$, $f'(1) \leq f(1)$ and $f'(2) \leq f(2)$. Thus, $f' \leq f$. Since no matter $f(1) \geq f(2)$ or $f(1) < f(2)$, $\max\{f'(1), f'(2)\} \geq f'(3)$, $V(f') = \min_{p \in D_2} E_p f'$. Since $f' \leq f$, $\min_{p \in D_2} E_p f' \leq \min_{p \in D_2} E_p f$. Hence, $V(f') \leq V(f)$. Let $g' \in \mathbb{R}^S$ be such that

$$g'(1) = \begin{cases} g(3) & \text{if } f(1) \geq f(2), \\ g(1) & \text{if } f(1) < f(2), \end{cases} \quad g'(2) = \begin{cases} g(2) & \text{if } f(1) \geq f(2), \\ g(3) & \text{if } f(1) < f(2), \end{cases} \quad g'(3) = g(3).$$

Since $\max\{g(1), g(2)\} < g(3)$, $g' \geq g$ and $\max\{g'(1), g'(2)\} = g'(3)$. Then $V(g') = \min_{p \in D_2} E_p g'$. Since $f \geq g$, $f' \geq g'$. Thus, $\min_{p \in D_2} E_p f' \geq \min_{p \in D_2} E_p g'$. Hence, $V(f') \geq V(g')$. Consider the sequence $\{\frac{1}{n} g + \frac{n-1}{n} g'\}_{n=1}^{\infty}$ of elements of $\mathbb{R}^S$. Since $\max\{g(1), g(2)\} < g(3)$, for each $n \in \mathbb{N}$, $\max\{\frac{1}{n} g(1) + \frac{n-1}{n} g'(1), \frac{1}{n} g(2) +$

$\frac{n-1}{n}g'(2)\} < \frac{1}{n}g(3) + \frac{n-1}{n}g'(3)$, and thus, $V(\frac{1}{n}g + \frac{n-1}{n}g') = \min_{p \in D_1} E_p(\frac{1}{n}g + \frac{n-1}{n}g')$. Since V is continuous, $V(g') = \lim_{n\to\infty} V(\frac{1}{n}g + \frac{n-1}{n}g') = \lim_{n\to\infty} \min_{p\in D_1} E_p(\frac{1}{n}g + \frac{n-1}{n}g') = \min_{p\in D_1} E_p g'$. Since $g' \geq g$, $\min_{p\in D_1} E_p g' \geq \min_{p\in D_1} E_p g$. Thus, $V(g') \geq V(g)$. Hence, $V(f) \geq V(f') \geq V(g') \geq V(g)$.

Now we prove that $\succsim$ does not admit a canonical WM representation. Suppose to the contrary that it admits a canonical WM representation $\langle u, D, \lambda\rangle$. Let $D_3 := \{p \in \Delta : p_3 \geq \frac{1}{4}\}$.

We first show that for each $f \in \mathbb{R}^S$, $V(f) \in [\min_{p\in D_2} E_p f, \max_{p\in D_2} E_p f]$. Let $f \in \mathbb{R}^S$. Suppose that $f(1) \leq f(2) < f(3)$. Then $V(f) = \min_{p\in D_1} E_p f = E_{p'} f$. Since $f(1) \leq f(2) < f(3)$, and since $p'_2 < p^*_2$ and $p'_3 < p^*_3$, $E_{p'} f \leq E_{p^*} f$. Recall $p'' = (\frac{2}{3}, \frac{1}{6}, \frac{1}{6})$. Since $f(1) \leq f(2) < f(3)$, and since $p'_1 = p''_1$ and $p'_3 > p''_3$, $E_{p'} f \geq E_{p''} f$. Thus, $V(f) \in [E_{p''} f, E_{p^*} f]$. Since $p'', p^* \in D_2$, $[E_{p''} f, E_{p^*} f] \subseteq [\min_{p\in D_2} E_p f, \max_{p\in D_2} E_p f]$. Hence, $V(f) \in [\min_{p\in D_2} E_p f, \max_{p\in D_2} E_p f]$. If $f(2) < f(1) < f(3)$, by analogous arguments, $V(f) \in [\min_{p\in D_2} E_p f, \max_{p\in D_2} E_p f]$. Lastly, if $\max\{f(1), f(2)\} \geq f(3)$, then $V(f) = \min_{p\in D_2} E_p f$, and thus, $V(f) \in [\min_{p\in D_2} E_p f, \max_{p\in D_2} E_p f]$.

We then show that for each $f \in \mathbb{R}^S$, $V(f) \in [\min_{p\in D_3} E_p f, \max_{p\in D_3} E_p f]$. Let $f \in \mathbb{R}^S$. Suppose that $\max\{f(1), f(2)\} < f(3)$. Then $V(f) = \min_{p\in D_1} E_p f$. Since $D_1 \subseteq D_3$, $\min_{p\in D_1} E_p f \in [\min_{p\in D_3} E_p f, \max_{p\in D_3} E_p f]$. Thus, $V(f) \in [\min_{p\in D_3} E_p f, \max_{p\in D_3} E_p f]$. Suppose that $\max\{f(1), f(2)\} \geq f(3)$. Then $V(f) = \min_{p\in D_2} E_p f$. Since $p^* \in D_2 \cap D_3$, $\min_{p\in D_2} E_p f \leq E_{p^*} f \leq \max_{p\in D_3} E_p f$. Thus, $V(f) \leq \max_{p\in D_3} E_p f$. To show that $V(f) \geq \min_{p\in D_3} E_p f$, we further consider the following three cases. First, suppose that $f(1) \leq f(3) \leq f(2)$. Recall that for each $p \in D_2$, by (29), $p_1 \leq \frac{2}{3}$. Since $f(1) \leq f(3) \leq f(2)$, and since $(\frac{3}{4}, 0, \frac{1}{4}) \in D_3$ and for each $p \in D_2$, $p_1 \leq \frac{2}{3} < \frac{3}{4}$, $\min_{p\in D_2} E_p f \geq \frac{3}{4} f(1) + \frac{1}{4} f(3) \geq \min_{p\in D_3} E_p f$. Thus, $V(f) \geq \min_{p\in D_3} E_p f$. Second, if $f(2) \leq f(3) \leq f(1)$, by analogous arguments, $V(f) \geq \min_{p\in D_3} E_p f$. Lastly, if $f(3) < \min\{f(1), f(2)\}$, then $\min_{p\in D_2} E_p f \geq f(3) = \min_{p\in D_3} E_p f$, and thus, $V(f) \geq \min_{p\in D_3} E_p f$.

We claim that $D \subseteq D_2 \cap D_3$. To see this, define $\lambda_2 : \mathbb{R}^S \to \mathbb{R}$ by setting for each $f \in \mathbb{R}^S$,

$$\lambda_2(f) := \begin{cases} \dfrac{\max_{p\in D_2} E_p f - V(f)}{\max_{p\in D_2} E_p f - \min_{p\in D_2} E_p f} & \text{if } \min_{p\in D_2} E_p f \neq \max_{p\in D_2} E_p f, \\ 1 & \text{if } \min_{p\in D_2} E_p f = \max_{p\in D_2} E_p f. \end{cases}$$

Thus, for each $f \in \mathbb{R}^S$, $V(f) = \lambda_2(f) \min_{p \in D_2} E_p f + (1 - \lambda_2(f)) \max_{p \in D_2} E_p f$, and since $V(f) \in [\min_{p \in D_2} E_p f, \max_{p \in D_2} E_p f]$, $\lambda_2(f) \in [0, 1]$. Since V is continuous, λ_2 is continuous on $\{f \in \mathbb{R}^S : \min_{p \in D_2} E_p f \neq \max_{p \in D_2} E_p f\}$. Let v be the identity mapping on $\mathbb{R}$. It can be readily seen that $\langle v, D_2, \lambda_2 \rangle$ is a weighed maxmin representation of $\succsim$. Analogously, since $V(f) \in [\min_{p \in D_3} E_p f, \max_{p \in D_3} E_p f]$, one can define $\lambda_3 : \mathbb{R}^S \to [0, 1]$ so that $\langle v, D_3, \lambda_3 \rangle$ is also a WM representation of $\succsim$. Since $\langle u, D, \lambda \rangle$ is a canonical WM representation of $\succsim$, $D \subseteq D_2 \cap D_3$.

Finally, to derive a contradiction, fix $g \in \mathbb{R}^S$ such that $g(1) < g(2) = g(3)$. Then $V(g) = \min_{p \in D_2} E_p g$. Recall that for each $p \in D_2$, by (29), $p_1 \leq \frac{2}{3}$, and it can also be readily seen that $p_1 = \frac{2}{3}$ if only if $p_2 = p_3 = \frac{1}{6}$. Thus, $V(g) = E_{p''} g$, and for each $p \in D_2 \backslash \{p''\}$, $E_p g > V(g)$. Since $p'' \notin D_3$ and $D \subseteq D_2 \cap D_3$, $D \subseteq D_2 \backslash \{p''\}$. Thus, $\min_{p \in D} E_p g > V(g) = v(x_g)$. Since both v and u are affine functions representing $\succsim$ restricted to X, by routine arguments, u is a positive affine transformation of v. Then $\min_{p \in D} E_p u(g) > u(x_g)$, and thus, $\lambda(u(g)) \min_{p \in D} E_p u(g) + (1 - \lambda(u(g))) \max_{p \in D} E_p u(g) > u(x_g)$. Since $\langle u, D, \lambda \rangle$ is a WM representation of $\succsim$, $g \succ x_g$, which is not possible. Hence, there is no canonical WM representation of $\succsim$. □

Proposition 9 *The preference relation $\succsim$ in Example 3 admits both a Bewley weighted maxmin representation and a canonical weighted maxmin representation, whereas the Bewley set of priors for $\succsim$ is not the smallest admissible set.*

Proof We first prove that $\succsim$ admits a Bewley WM representation. By Proposition 1, it is equivalent to prove that $\succsim$ satisfies A.1–A.4 and A.6. Clearly, it satisfies A.1. Since for each $t \in \mathbb{R}$, $V(t\mathbf{1}) = t$, it satisfies A.2 and A.6.

To show that $\succsim$ satisfies A.3, it suffices to show the continuity of V. Let $f \in \mathbb{R}^S$ and $\{f^n\}_{n=1}^{\infty}$ be a sequence of elements in $\mathbb{R}^S$ that converges to f. Before showing that $\lim_{n\to\infty} V(f^n) = V(f)$, we assume first that $\max\{f(1), f(2)\} < f(3)$ and for each $n \in \mathbb{N}$, $\max\{f^n(1), f^n(2)\} < f^n(3)$, and check that $\lim_{n\to\infty} \alpha(f^n) = \alpha(f)$. Suppose that $f(1) < f(2) < f(3)$. Then $\alpha(f) = \text{med}\{0, \frac{1-E_{p''}f}{1-E_{p''}f+E_{p'}f}, 1\}$. Moreover, for sufficiently large n, $f^n(1) < f^n(2) < f^n(3)$, so that $\alpha(f^n) = \text{med}\{0, \frac{1-E_{p''}f^n}{1-E_{p''}f^n+E_{p'}f^n}, 1\}$. Since the median operator is continuous, $\lim_{n\to\infty} \alpha(f^n) = \alpha(f)$. Suppose that $f(2) < f(1) < f(3)$. By analogous arguments, $\lim_{n\to\infty} \alpha(f^n) = \alpha(f)$. Suppose that $f(1) = f(2) < f(3)$. Since $f(1) = f(2) < f(3)$, and since $p_3' = q_3'$ and $p_3'' = q_3''$, $\frac{1-E_{p''}f}{1-E_{p''}f+E_{p'}f} = \frac{1-E_{q''}f}{1-E_{q''}f+E_{q'}f}$. Then $\alpha(f) = \lim_{n\to\infty} \text{med}\{0, \frac{1-E_{p''}f^n}{1-E_{p''}f^n+E_{p'}f^n}, 1\} = \lim_{n\to\infty} \text{med}\{0, \frac{1-E_{q''}f^n}{1-E_{q''}f^n+E_{q'}f^n}, 1\}$. For each $n \in \mathbb{N}$, either $\alpha(f^n) = \text{med}\{0, \frac{1-E_{p''}f^n}{1-E_{p''}f^n+E_{p'}f^n}, 1\}$ or $\alpha(f^n) = \text{med}\{0, \frac{1-E_{q''}f^n}{1-E_{q''}f^n+E_{q'}f^n}, 1\}$. Thus, $\lim_{n\to\infty} \alpha(f^n) = \alpha(f)$. We now show that $\lim_{n\to\infty} V(f^n) = V(f)$ in each of the following two cases.

Case 1: $\max\{f(1), f(2)\} \neq f(3)$. Suppose that $\max\{f(1), f(2)\} < f(3)$. Then $V(f) = \min_{p \in D_1(f)} E_p f$. Moreover, for sufficiently large n, $\max\{f^n(1), f^n(2)\} < f^n(3)$, so that $V(f^n) = \min_{p \in D_1(f^n)} E_p f^n$. Since α is continuous at f, D_1 is continuous (i.e., both upper and lower hemicontinuous) at f. Thus, by the maximum theorem, $\lim_{n\to\infty} V(f^n) = V(f)$. Suppose that $\max\{f(1), f(2)\} > f(3)$. Then $V(f) = \min_{p \in D_2} E_p f$. Moreover, for sufficiently large n, $\max\{f^n(1), f^n(2)\} > f^n(3)$, so that $V(f^n) = \min_{p \in D_2} E_p f^n$. Thus, $\lim_{n\to\infty} V(f^n) = V(f)$.

Case 2: $\max\{f(1), f(2)\} = f(3)$. Then $V(f) = \min_{p \in D_2} E_p f$. Suppose that $f(1) < f(2)$. Thus, $f(1) < f(2) = f(3)$ and for sufficiently large n, $f^n(1) < f^n(2)$. Recall that as shown in the proof of Proposition 8, for each $p \in D_2$, by (29), $p_1 \leq \frac{2}{3} = p''_1$, and thus, $\min_{p \in D_2} E_p f = E_{p''} f = E_{p'} f$. Then

$$\lim_{n\to\infty} E_{p'} f^n = \lim_{n\to\infty} E_{p''} f^n = \lim_{n\to\infty} \min_{p \in D_2} E_p f^n = V(f). \tag{32}$$

For sufficiently large n, since $f^n(1) < f^n(2)$, and since $p'_1 > q'_1$, $p''_1 > q''_1$, $p'_3 = q'_3$, and $p''_3 = q''_3$, for each $\beta \in [0, 1]$, $E_{\beta p' + (1-\beta) p''} f^n < E_{\beta q' + (1-\beta) q''} f^n$, and thus,

$$\begin{aligned} &\text{either } V(f^n) = \min_{p \in D_1(f^n)} E_p f^n = \alpha(f^n) E_{p'} f^n + (1 - \alpha(f^n)) E_{p''} f^n \\ &\quad \text{or } V(f^n) = \min_{p \in D_2} E_p f^n. \end{aligned} \tag{33}$$

By (32) and (33), $\lim_{n\to\infty} V(f^n) = V(f)$. Similarly, one can show that when $f(1) > f(2)$, $\lim_{n\to\infty} V(f^n) = V(f)$. Suppose that $f(1) = f(2)$. Then $f(1) = f(2) = f(3)$, and thus, for each $p \in \Delta$, $\lim_{n\to\infty} E_p f^n = \lim_{n\to\infty} \min_{p \in D_2} E_p f^n = \min_{p \in D_2} E_p f = V(f)$. For each $n \in \mathbb{N}$, either $V(f^n) = E_{\alpha(f^n) p' + (1-\alpha(f^n)) p''} f^n$, or $V(f^n) = E_{\alpha(f^n) q' + (1-\alpha(f^n)) q''} f^n$, or $V(f^n) = \min_{p \in D_2} E_p f^n$. Thus, $\lim_{n\to\infty} V(f^n) = V(f)$.

Lastly, we show that $\succsim$ satisfies A.4. To facilitate our proof, we first derive an equivalent expression of V. Let $f \in \mathbb{R}^S$. Suppose that $f(1) \leq f(2) < f(3)$. Since $f(1) \leq f(2) < f(3)$, and since $p'_1 = p''_1$ and $p'_3 > p''_3$, $E_{p'} f > E_{p''} f$. Thus,

$$\alpha(f) = \begin{cases} 1 & \text{if } E_{p'} f \leq 0, \\ \dfrac{1 - E_{p''} f}{1 - E_{p''} f + E_{p'} f} & \text{if } E_{p'} f > 0 \text{ and } E_{p''} f < 1, \\ 0 & \text{if } E_{p''} f \geq 1. \end{cases}$$

Moreover, since $f(1) \leq f(2) < f(3)$, and since $p'_1 > q'_1$, $p''_1 > q''_1$, $p'_3 = q'_3$, and $p''_3 = q''_3$, for each $\beta \in [0, 1]$, $E_{\beta p' + (1-\beta) p''} f < E_{\beta q' + (1-\beta) q''} f$. Thus,

$$V(f) = \min_{p \in D_1(f)} E_p f = E_{\alpha(f) p' + (1-\alpha(f)) p''} f = \alpha(f) E_{p'} f + (1 - \alpha(f)) E_{p''} f. \tag{34}$$

Similarly, if $f(2) < f(1) < f(3)$, then

$$\alpha(f) = \begin{cases} 1 & \text{if } E_{q'} f \leq 0, \\ \dfrac{1 - E_{q''} f}{1 - E_{q''} f + E_{q'} f} & \text{if } E_{q'} f > 0 \text{ and } E_{q''} f < 1, \\ 0 & \text{if } E_{q''} f \geq 1, \end{cases}$$

and $V(f) = \alpha(f) E_{q'} f + (1 - \alpha(f)) E_{q''} f$.

Hence, for each $f \in \mathbb{R}^S$,

$$V(f) = \begin{cases} E_{p'} f & \text{if } f(1) \leq f(2) < f(3) \text{ and } E_{p'} f \leq 0, \\ \dfrac{E_{p'} f}{1 - E_{p''} f + E_{p'} f} & \text{if } f(1) \leq f(2) < f(3),\ E_{p'} f > 0, \text{ and } E_{p''} f < 1, \\ E_{p''} f & \text{if } f(1) \leq f(2) < f(3) \text{ and } E_{p''} f \geq 1, \\ E_{q'} f & \text{if } f(2) < f(1) < f(3) \text{ and } E_{q'} f \leq 0, \\ \dfrac{E_{q'} f}{1 - E_{q''} f + E_{q'} f} & \text{if } f(2) < f(1) < f(3),\ E_{q'} f > 0, \text{ and } E_{q''} f < 1, \\ E_{q''} f & \text{if } f(2) < f(1) < f(3) \text{ and } E_{q''} f \geq 1, \\ \min\limits_{p \in D_2} E_p f & \max\{f(1), f(2)\} \geq f(3). \end{cases}$$

Now let $f, g \in \mathbb{R}^S$ be such that $f \geq g$. We shall show that $V(f) \geq V(g)$ in each of the following four cases.

Case 1: Either $f(1) \leq f(2) < f(3)$ and $g(1) \leq g(2) < g(3)$, or $f(2) < f(1) < f(3)$ and $g(2) < g(1) < g(3)$. Consider the former scenario first. Suppose that $E_{p'} f \leq 0$. Then $V(f) = E_{p'} f$ and $E_{p'} g \leq E_{p'} f \leq 0$. Thus, $V(g) = E_{p'} g \leq V(f)$. Suppose that $E_{p'} f > 0$ and $E_{p''} f < 1$. Then $V(f) = \frac{E_{p'} f}{1 - E_{p''} f + E_{p'} f} > 0$ and $E_{p''} g \leq E_{p''} f < 1$. If $E_{p'} g \leq 0$, then $V(g) = E_{p'} g \leq 0 < V(f)$. If $E_{p'} g > 0$, then $V(g) = \frac{E_{p'} g}{1 - E_{p''} g + E_{p'} g} \leq \frac{E_{p'} f}{1 - E_{p''} f + E_{p'} f} = V(f)$, where the inequality holds since $0 < E_{p'} g \leq E_{p'} f$ and $0 < 1 - E_{p''} f \leq 1 - E_{p''} g$. Suppose that $E_{p''} f \geq 1$. Then $V(f) = E_{p''} f \geq 1$. If $E_{p'} g \leq 0$, then $V(g) = E_{p'} g \leq 0 < 1 \leq V(f)$. If $E_{p'} g > 0$ and $E_{p''} g < 1$, then $V(g) = \frac{E_{p'} g}{1 - E_{p''} g + E_{p'} g} < 1 \leq V(f)$. If $E_{p''} g \geq 1$, then $V(g) = E_{p''} g \leq E_{p''} f = V(f)$. In the latter scenario, by analogous arguments, $V(f) \geq V(g)$.

Case 2: $\max\{f(1), f(2)\} \geq f(3)$ and $\max\{g(1), g(2)\} \geq g(3)$. Then $V(f) = \min\limits_{p \in D_2} E_p f \geq \min\limits_{p \in D_2} E_p g = V(g)$.

Case 3: Either $f(1) \leq f(2) < f(3)$ and $g(2) < g(1) < g(3)$, or $f(2) < f(1) < f(3)$ and $g(1) \leq g(2) < g(3)$. Consider the former scenario first. Let $f' \in \mathbb{R}^S$ be such that

$$f'(1) = f(1), \quad f'(2) = f(1), \quad f'(3) = f(3).$$

Then $f \geq f' \geq g$. Since $f \geq f'$, and since $f(1) \leq f(2) < f(3)$ and $f'(1) = f'(2) < f'(3)$, by the result of Case 1, $V(f) \geq V(f')$. Consider the sequence $\{\frac{n-1}{n} f' + \frac{1}{n} g\}_{n=1}^{\infty}$ of elements in $\mathbb{R}^S$. For each $n \in \mathbb{N}$, since $f' \geq g$, $\frac{n-1}{n} f' + \frac{1}{n} g \geq g$; since $f'(2) = f'(1) < f'(3)$ and $g(2) < g(1) < g(3)$, $\frac{n-1}{n} f'(2) + \frac{1}{n} g(2) < \frac{n-1}{n} f'(1) + \frac{1}{n} g(1) < \frac{n-1}{n} f'(3) + \frac{1}{n} g(3)$. For each $n \in \mathbb{N}$, since $\frac{n-1}{n} f' + \frac{1}{n} g \geq g$, and since $\frac{n-1}{n} f'(2) + \frac{1}{n} g(2) < \frac{n-1}{n} f'(1) + \frac{1}{n} g(1) < \frac{n-1}{n} f'(3) + \frac{1}{n} g(3)$ and $g(2) < g(1) < g(3)$, by the result of Case 1, $V(\frac{n-1}{n} f' + \frac{1}{n} g) \geq V(g)$. Since V is continuous (as shown when proving that A.3 holds), $V(f') = \lim_{n \to \infty} V(\frac{n-1}{n} f' + \frac{1}{n} g)$. Thus, $V(f') \geq V(g)$. Hence, $V(f) \geq V(f') \geq V(g)$. In the latter scenario, by analogous arguments, $V(f) \geq V(g)$.

Case 4: Either $\max\{f(1), f(2)\} < f(3)$ and $\max\{g(1), g(2)\} \geq g(3)$, or $\max\{f(1), f(2)\} \geq f(3)$ and $\max\{g(1), g(2)\} < g(3)$. Consider the former scenario first. Let $f' \in \mathbb{R}^S$ be such that

$$f'(1) = f(1), \quad f'(2) = f(2), \quad f'(3) = \max\{f(1), f(2)\}.$$

Then $f \geq f' \geq g$ and $\max\{f'(1), f'(2)\} = f'(3)$. Since $f' \geq g$, and since $\max\{f'(1), f'(2)\} = f'(3)$ and $\max\{g(1), g(2)\} \geq g(3)$, by the result of Case 2, $V(f') \geq V(g)$. Consider the sequence $\{\frac{n-1}{n} f' + \frac{1}{n} f\}_{n=1}^{\infty}$ of elements in $\mathbb{R}^S$. For each $n \in \mathbb{N}$, since $f \geq f'$, $f \geq \frac{n-1}{n} f' + \frac{1}{n} f$; since $f'(3) = \max\{f(1), f(2)\} < f(3)$, $\max\{\frac{n-1}{n} f'(1) + \frac{1}{n} f(1), \frac{n-1}{n} f'(2) + \frac{1}{n} f(2)\} = \max\{f(1), f(2)\} < \frac{n-1}{n} f'(3) + \frac{1}{n} f(3)$. For each $n \in \mathbb{N}$, since $f \geq \frac{n-1}{n} f' + \frac{1}{n} f$, and since $\max\{f(1), f(2)\} < f(3)$ and $\max\{\frac{n-1}{n} f'(1) + \frac{1}{n} f(1), \frac{n-1}{n} f'(2) + \frac{1}{n} f(2)\} < \frac{n-1}{n} f'(3) + \frac{1}{n} f(3)$, by the results of Cases 1 and 3, $V(f) \geq V(\frac{n-1}{n} f' + \frac{1}{n} f)$. Since V is continuous, $V(f') = \lim_{n \to \infty} V(\frac{n-1}{n} f' + \frac{1}{n} f)$. Thus, $V(f) \geq V(f')$. Hence, $V(f) \geq V(f') \geq V(g)$. In the latter scenario, by similar arguments, $V(f) \geq V(g)$.

Now we prove that $\succsim$ admits a canonical WM representation with $D := D_2 \cap \{p \in \Delta : p_3 \geq \frac{1}{6}\}$ being the smallest admissible set. Let $f \in \mathbb{R}^S$. We first show that $V(f) \in [\min_{p \in D} E_p f, \max_{p \in D} E_p f]$ in each of the following two cases.

Case 1: Either $f(1) \leq f(2) < f(3)$ or $f(2) < f(1) < f(3)$. Consider the former scenario first. Then $E_{p''} f < E_{p'} f < E_{p^*} f$, and as shown in (34), $V(f) = \alpha(f) E_{p'} f + (1 - \alpha(f)) E_{p''} f$. Thus, $V(f) \in [E_{p''} f, E_{p^*} f]$. Since $V(f) \in [E_{p''} f, E_{p^*} f]$ and $p'', p^* \in D$, $V(f) \in [\min_{p \in D} E_p f, \max_{p \in D} E_p f]$. In the latter scenario, by analogous arguments, $V(f) \in [\min_{p \in D} E_p f, \max_{p \in D} E_p f]$.

Case 2: Either $f(1) \leq f(2)$ and $f(3) \leq f(2)$, or $f(1) > f(2)$ and $f(3) \leq f(1)$. Consider the former scenario first. Then $V(f) = \min_{p \in D_2} E_p f$. Since $p^* \in D_2$ and $p^* \in D$, $V(f) \leq E_{p^*} f \leq \max_{p \in D} E_p f$. To show that $V(f) \geq \min_{p \in D} E_p f$, suppose to the contrary that $V(f) < \min_{p \in D} E_p f$. Then f is not a constant act. Let $\bar{p} \in \arg\min_{p \in D_2} E_p f$. Then $E_{\bar{p}} f = V(f) < \min_{p \in D} E_p f$. Thus, $\bar{p} \notin D$, i.e., $\bar{p}_3 < \frac{1}{6}$. Recall that as argued in the proof of Proposition 8, for each $p \in D_2$, by (29), $p_1 \leq \frac{2}{3}$, and $p_1 = \frac{2}{3}$ if and only if $p_2 = p_3 = \frac{1}{6}$. Since $\bar{p} \in D_2$ and $\bar{p}_3 < \frac{1}{6}$, $\bar{p}_1 < \frac{2}{3}$. Since $f(1), f(3) \leq f(2)$ and f is not a constant act, and since $\bar{p}_1 < \frac{2}{3} = p''_1$ and $\bar{p}_3 < \frac{1}{6} = p''_3$, $E_{\bar{p}} f > E_{p''} f$. Since $p'' \in D_2$, $E_{p''} f \geq \min_{p \in D_2} E_p f$. Thus, $E_{\bar{p}} f > \min_{p \in D_2} E_p f$, which contradicts $\bar{p} \in \arg\min_{p \in D_2} E_p f$. In the latter scenario, by analogous arguments, $V(f) \in [\min_{p \in D} E_p f, \max_{p \in D} E_p f]$.

Next, we show that $\succsim$ admits a WM representation with D being an admissible set. Since for each $f \in \mathbb{R}^S$, $V(f) \in [\min_{p \in D} E_p f, \max_{p \in D} E_p f]$, there is $\lambda : \mathbb{R}^S \to [0, 1]$ such that for each $f \in \mathbb{R}^S$, $V(f) = \lambda(f) \min_{p \in D} E_p f + (1 - \lambda(f)) \max_{p \in D} E_p f$. Since V is continuous, λ is continuous on $\{f \in \mathbb{R}^S : \min_{p \in D} E_p f \neq \max_{p \in D} E_p f\}$. Let u be the identity mapping on $\mathbb{R}$. Then $\langle u, D, \lambda \rangle$ is a WM representation of $\succsim$.

We then show that D is actually the smallest admissible set, so that $\langle u, D, \lambda \rangle$ is a canonical WM representation of $\succsim$. Let $\langle u', D', \lambda' \rangle$ be another WM representation of $\succsim$. To show that $D \subseteq D'$, suppose to the contrary that there is $\bar{p} \in D \backslash D'$. By a standard separation theorem, there is a non-constant act $\bar{f} \in \mathbb{R}^S$ such that $E_{\bar{p}} \bar{f} < \min_{p \in D'} E_p \bar{f}$. Let $t \in \mathbb{R}$ be such that $E_{p''}(\bar{f} + t\mathbf{1}) \geq 1$ and $E_{q''}(\bar{f} + t\mathbf{1}) \geq 1$. We claim that $V(\bar{f}+t\mathbf{1}) \leq E_{\bar{p}}(\bar{f}+t\mathbf{1})$. To see this, suppose first $\max\{\bar{f}(1)+t, \bar{f}(2)+t\} \geq \bar{f}(3)+t$. Then $V(\bar{f} + t\mathbf{1}) = \min_{p \in D_2} E_p(\bar{f} + t\mathbf{1})$. Since $\bar{p} \in D$ and $D \subseteq D_2$, $\min_{p \in D_2} E_p(\bar{f} + t\mathbf{1}) \leq E_{\bar{p}}(\bar{f}+t\mathbf{1})$. Thus, $V(\bar{f}+t\mathbf{1}) \leq E_{\bar{p}}(\bar{f}+t\mathbf{1})$. Suppose next that $\bar{f}(1)+t \leq \bar{f}(2)+t < \bar{f}(3)+t$. Since $\bar{f}(1)+t \leq \bar{f}(2)+t < \bar{f}(3)+t$ and $E_{p''}(\bar{f}+t\mathbf{1}) \geq 1$, $V(\bar{f}+t\mathbf{1}) = E_{p''}(\bar{f} + t\mathbf{1})$. Recall that as argued in the proof of Proposition 8, for each $p \in D_2$, $p_1 \leq \frac{2}{3}$. Since $\bar{p} \in D_2$, $\bar{p}_1 \leq \frac{2}{3} = p''_1$. Since $\bar{p} \in D$, $\bar{p}_3 \geq \frac{1}{6} = p''_3$. Since $\bar{f}(1)+t \leq \bar{f}(2)+t < \bar{f}(3)+t$, and since $\bar{p}_1 \leq p''_1$ and $\bar{p}_3 \geq p''_3$, $E_{p''}(\bar{f} + t\mathbf{1}) \leq E_{\bar{p}}(\bar{f} + t\mathbf{1})$. Thus, $V(\bar{f}+t\mathbf{1}) \leq E_{\bar{p}}(\bar{f}+t\mathbf{1})$. Lastly, if $\bar{f}(2)+t < \bar{f}(1)+t < \bar{f}(3)+t$, by analogous arguments as in the case $\bar{f}(1)+t \leq \bar{f}(2)+t < \bar{f}(3)+t$, $V(\bar{f}+t\mathbf{1}) \leq E_{\bar{p}}(\bar{f}+t\mathbf{1})$.

Since $V(\bar{f}+t\mathbf{1}) \leq E_{\bar{p}}(\bar{f}+t\mathbf{1})$ and $E_{\bar{p}} \bar{f} < \min_{p \in D'} E_p \bar{f}$, $V(\bar{f}+t\mathbf{1}) < \min_{p \in D'} E_p(\bar{f} + t\mathbf{1})$. Recall that u is the identity mapping on $\mathbb{R}$. Thus, $u(x_{\bar{f}+t\mathbf{1}}) < \min_{p \in D'} E_p(\bar{f} + t\mathbf{1})$. Since both u and u' are affine functions representing $\succsim$ restricted to X, by routine arguments, u' is a positive affine transformation of u. Then $u'(x_{\bar{f}+t\mathbf{1}}) < \min_{p \in D'} E_p u'(\bar{f} + t\mathbf{1}) \leq \lambda'(u'(\bar{f} + t\mathbf{1})) \min_{p \in D'} E_p u'(\bar{f} + t\mathbf{1}) + (1 - \lambda'(u'(\bar{f} + t\mathbf{1}))) \max_{p \in D'} E_p u'(\bar{f} + t\mathbf{1})$. Since $\langle u', D', \lambda' \rangle$ is a WM representation of $\succsim$, $\bar{f} + t\mathbf{1} \succ x_{\bar{f}+t\mathbf{1}}$, which is not possible. Hence, $D \subseteq D'$, as desired.

Finally, we show that D is not the Bewley set. Suppose to the contrary that D is the Bewley set. That is, there is a WM representation $\langle u'', D, \lambda'' \rangle$ of $\succsim$ such that $\langle u'', D \rangle$ is a Bewley representation of the unambiguous preference relation $\succsim^*$ induced from $\succsim$. Then u'' is a positive affine transformation of the identity mapping u. Observe that $p' \notin D_2$. Thus, $p' \notin D$. By a standard separation theorem, there is a non-constant act $f' \in \mathbb{R}^S$ such that $E_{p'} f' < \min_{p \in D} E_p f'$. Let

$x := \min_{p \in D} E_p f'$. Thus, $E_{p'} f' < x$ and $\min_{p \in D} E_p u''(f') = u''(x)$. Let $g \in \mathbb{R}^S$ be such that $g(1) < g(2) < g(3)$, $f'(1) + g(1) < f'(2) + g(2) < f'(3) + g(3)$, and $E_{p'}(x\mathbf{1} + g) \leq 0$. Since $E_{p'} f' < x$, $E_{p'}(\frac{1}{2} f' + \frac{1}{2} g) < E_{p'}(\frac{1}{2} x\mathbf{1} + \frac{1}{2} g) \leq 0$. Then $V(\frac{1}{2} f' + \frac{1}{2} g) = E_{p'}(\frac{1}{2} f' + \frac{1}{2} g) < E_{p'}(\frac{1}{2} x\mathbf{1} + \frac{1}{2} g) = V(\frac{1}{2} x + \frac{1}{2} g)$. Thus, $\frac{1}{2} x + \frac{1}{2} g \succ \frac{1}{2} f' + \frac{1}{2} g$. Since $\langle u'', D\rangle$ is a Bewley representation of $\succsim^*$, and since for each $p \in D$, $E_p u''(f') \geq \min_{p \in D} E_p u''(f') = u''(x)$, $f' \succsim^* x$. Thus, $\frac{1}{2} f' + \frac{1}{2} g \succsim \frac{1}{2} x + \frac{1}{2} g$, which contradicts $\frac{1}{2} x + \frac{1}{2} g \succ \frac{1}{2} f' + \frac{1}{2} g$. Hence, D is not the Bewley set. □

Lemma 2 *Let $\varphi \in \mathbb{R}^S$, $t \in [\varphi_*, \varphi^*]$, and $B := \{p \in \Delta : E_p \varphi = t\}$. Let K be a non-empty closed subset of Δ such that for each $p \in K$, $E_p \varphi \geq t$. Let $c := \min_{p \in B} d(p, K)$. Then $\min_{p \in \Delta : d(p,K) \leq c} E_p \varphi = t$.*

Proof Let $\bar{p} \in B$ be such that $d(\bar{p}, K) = c$. Since $\bar{p} \in B$, $E_{\bar{p}} \varphi = t$. Since $d(\bar{p}, K) = c$, $E_{\bar{p}} \varphi \geq \min_{p \in \Delta : d(p,K) \leq c} E_p \varphi$. Thus, $t \geq \min_{p \in \Delta : d(p,K) \leq c} E_p \varphi$. Suppose that $t > \min_{p \in \Delta : d(p,K) \leq c} E_p \varphi$. Let $p' \in \Delta$ and $q' \in K$ be such that $d(p', q') = d(p', K) \leq c$ and $E_{p'} \varphi = \min_{p \in \Delta : d(p,K) \leq c} E_p \varphi$. Then $t > E_{p'} \varphi$. Since $q' \in K$ and for each $p \in K$, $E_p \varphi \geq t$, $E_{q'} \varphi \geq t$. Since $E_{q'} \varphi \geq t > E_{p'} \varphi$, $p' \neq q'$ and there is $\alpha \in [0, 1)$ such that $t = E_{\alpha p' + (1-\alpha) q'} \varphi$. Thus, $\alpha p' + (1-\alpha) q' \in B$ and $d(\alpha p' + (1-\alpha) q', q') < d(p', q')$. Hence,

$$c = \min_{p \in B} d(p, K) \leq d(\alpha p' + (1-\alpha) q', q') < d(p', q') = d(p', K) \leq c,$$

which is not possible. Hence, $t = \min_{p \in \Delta : d(p,K) \leq c} E_p \varphi$. □

Proof of Theorem 2 To prove the sufficiency of the axioms, let $\succsim$ satisfy A.1–A.4, A.5.2, and A.6. By Lemma 1, there exist an affine onto function $u : X \to \mathbb{R}$ and a normalized, non-decreasing, and continuous functional $I : \mathbb{R}^S \to \mathbb{R}$ satisfying (22).

Let $K := \{p \in \Delta : \text{for each } \varphi \in \mathbb{R}^S, I(\varphi) \leq E_p \varphi\}$. Equivalently, K is the benchmark set. The equivalence has been shown by Ghirardato and Marinacci (2002) for biseparable preference relations (see their Theorem 12). It can be readily seen that their proof applies here. Since K is the benchmark set, by A.5.2, K is non-empty. Since I is continuous, K is closed. By the definition of K, it is convex.

Define $B : \mathbb{R}^S \rightrightarrows \Delta$ and $\sigma : \mathbb{R}^S \to \mathbb{R}_+$ by setting for each $\varphi \in \mathbb{R}^S$,

$$B(\varphi) := \{p \in \Delta : I(\varphi) = E_p \varphi\},$$

and

$$\sigma(\varphi) := \min_{p \in B(\varphi)} d(p, K). \tag{35}$$

For each $\varphi \in \mathbb{R}^S$, it can be readily seen that $B(\varphi)$ is closed, and since I is normalized and non-decreasing, $I(\varphi) \in [\varphi_*, \varphi^*]$, so that $B(\varphi)$ is non-empty. Thus, σ is well-defined. Since I is non-decreasing, $\sigma \in \Sigma(K)$. Moreover, by Lemma 2, for each $\varphi \in \mathbb{R}^S$, $I(\varphi) = \min_{p \in \Delta : d(p,K) \leq \sigma(\varphi)} E_p\varphi$. Thus, by (22), for each pair $f, g \in \mathcal{F}$,

$$f \succsim g \iff \min_{p \in \Delta : d(p,K) \leq \sigma(u(f))} E_p u(f) \geq \min_{p \in \Delta : d(p,K) \leq \sigma(u(g))} E_p u(g),$$

Therefore, if we can show that $\sigma : \mathbb{R}^S \to \mathbb{R}_+$ is continuous on $\{u(f) \in \mathbb{R}^S : f \in \mathcal{F}, f \nsim x_{*f}\}$ and lower semicontinuous on $\{u(f) \in \mathbb{R}^S : f \in \mathcal{F}, f \sim x_{*f}\}$, then $\langle u, K, \sigma \rangle$ is a VC representation of $\succsim$.

To show the continuity property of σ, let $\varphi \in \mathbb{R}^S$ and $f \in \mathcal{F}$ be such that $u(f) = \varphi$, and let $\{\varphi^n\}_{n=1}^{\infty}$ be a sequence of elements in $\mathbb{R}^S$ such that $\lim_{n\to\infty} \varphi^n = \varphi$. Suppose first that $f \nsim x_{*f}$. Then by (22), $I(u(f)) \neq I(u(x_{*f}))$, and thus, $I(\varphi) \neq I(\varphi_* \mathbf{1})$. Since $I(\varphi) \neq I(\varphi_* \mathbf{1})$, and since I is normalized and non-decreasing, $\varphi_* < I(\varphi) \leq \varphi^*$. We shall show that $\lim_{n\to\infty} \sigma(\varphi^n) = \sigma(\varphi)$ in each of the following two cases.

Case 1: $I(\varphi) = \varphi^*$. Let $q^* \in K$. By the definition of K, $I(\varphi) \leq E_{q^*}\varphi$. Since $I(\varphi) \leq E_{q^*}\varphi \leq \varphi^* = I(\varphi)$, $E_{q^*}\varphi = I(\varphi)$. Thus, $q^* \in B(\varphi)$. Since $q^* \in K$ and $q^* \in B(\varphi)$, $\sigma(\varphi) = \min_{p \in B(\varphi)} d(p, K^*) = d(q^*, q^*) = 0$. To show that $\lim_{n\to\infty} \sigma(\varphi^n) = 0$, suppose to the contrary that there exist $\epsilon > 0$ and a subsequence $\{\varphi^{n_m}\}_{m=1}^{\infty}$ of $\{\varphi^n\}_{n=1}^{\infty}$ such that for each $m \in \mathbb{N}$, $\sigma(\varphi^{n_m}) > \epsilon$. Let $q_* \in \Delta$ be such that $E_{q_*}\varphi = \varphi_*$. Since $I(\varphi) > \varphi_*$, $I(\varphi) > E_{q_*}\varphi$. Thus, $q_* \notin K$. Let $\alpha \in (0, 1)$ be such that $d(\alpha q_* + (1-\alpha)q^*, q^*) < \epsilon$. Since $q^* \in K$, $d(\alpha q_* + (1-\alpha)q^*, K) \leq d(\alpha q_* + (1-\alpha)q^*, q^*) < \epsilon$. Then for each $m \in \mathbb{N}$, $d(\alpha q_* + (1-\alpha)q^*, K) < \sigma(\varphi^{n_m})$, and thus, $I(\varphi^{n_m}) = \min_{p \in \Delta : d(p,K) \leq \sigma(\varphi^{n_m})} E_p\varphi^{n_m} \leq E_{\alpha q_* + (1-\alpha)q^*}\varphi^{n_m}$. Since I is continuous, $I(\varphi) = \lim_{m\to\infty} I(\varphi^{n_m})$. Thus, $I(\varphi) = \lim_{m\to\infty} I(\varphi^{n_m}) \leq \lim_{m\to\infty} E_{\alpha q_* + (1-\alpha)q^*}\varphi^{n_m} = \alpha E_{q_*}\varphi + (1-\alpha)E_{q^*}\varphi = \alpha\varphi_* + (1-\alpha)I(\varphi) < I(\varphi)$, which contradicts $I(\varphi) = \varphi^*$, as desired.

Case 2: $\varphi_* < I(\varphi) < \varphi^*$. To show the continuity of σ at φ, by the maximum theorem, it suffices to show the continuity of the correspondence B at φ. To show the upper hemicontinuity of B at φ, let $\bar{p} \in \Delta$ and let $\{p^n\}_{n=1}^{\infty}$ be a sequence of elements in Δ such that $\lim_{n\to\infty} p^n = \bar{p}$ and for each $n \in \mathbb{N}$, $p^n \in B(\varphi^n)$. We shall show that $\bar{p} \in B(\varphi)$. For each $n \in \mathbb{N}$, since $p^n \in B(\varphi^n)$, $E_{p^n}\varphi^n = I(\varphi^n)$. Thus, $E_{\bar{p}}\varphi = \lim_{n\to\infty} E_{p^n}\varphi^n = \lim_{n\to\infty} I(\varphi^n)$, and since I is continuous, $E_{\bar{p}}\varphi = I(\varphi)$. Hence, $\bar{p} \in B(\varphi)$. To show the lower hemicontinuity of B at φ, let $\bar{p} \in B(\varphi)$. We shall show that there exist a sequence $\{p^m\}_{m=1}^{\infty}$ and a subsequence $\{\varphi^{n_m}\}_{m=1}^{\infty}$ of $\{\varphi^n\}_{n=1}^{\infty}$ such that $\lim_{m\to\infty} p^m = \bar{p}$ and for each $m \in \mathbb{N}$, $p^m \in B(\varphi^{n_m})$. Since $\bar{p} \in B(\varphi)$, $E_{\bar{p}}\varphi = I(\varphi)$. For each $\epsilon > 0$, let $A(\epsilon) := \{p \in \Delta : d(p, \bar{p}) \leq \epsilon\}$. For each $\epsilon > 0$, since $\varphi_* < E_{\bar{p}}\varphi = I(\varphi) < \varphi^*$, $\min_{p \in A(\epsilon)} E_p\varphi < I(\varphi) < \max_{p \in A(\epsilon)} E_p\varphi$, and since I is continuous, for sufficiently large n, $\min_{p \in A(\epsilon)} E_p\varphi^n < I(\varphi^n) < \max_{p \in A(\epsilon)} E_p\varphi^n$. Thus, for each $\epsilon > 0$, when n is sufficiently large, there is $p \in A(\epsilon)$ such that $E_p\varphi^n = I(\varphi^n)$, so that $p \in A(\epsilon) \cap B(\varphi^n)$. Thus, there exist a sequence $\{p^m\}_{m=1}^{\infty}$ and a subsequence

$\{\varphi^{n_m}\}_{m=1}^{\infty}$ of $\{\varphi^n\}_{n=1}^{\infty}$ such that for each $m \in \mathbb{N}$, $p^m \in A(\frac{1}{m}) \cap B(\varphi^{n_m})$. Since for each $m \in \mathbb{N}$, $p^m \in A(\frac{1}{m})$, $\lim_{m\to\infty} p^m = \bar{p}$.

Now suppose that $f \sim x_{*f}$. We want to show that $\liminf_{n\to\infty} \sigma(\varphi^n) \geq \sigma(\varphi)$. Since for each $n \in \mathbb{N}$, $\sigma(\varphi^n) \geq 0$, $\liminf_{n\to\infty} \sigma(\varphi^n) \geq 0$. If φ is constant, then $B(\varphi) = \Delta$, and thus, $\sigma(\varphi) = \min_{p\in B(\varphi)} d(p, K^*) = 0 \leq \liminf_{n\to\infty} \sigma(\varphi^n)$. Suppose that φ is not constant. Suppose to the contrary that $\liminf_{n\to\infty} \sigma(\varphi^n) < \sigma(\varphi)$. Then there is a convergent subsequence $\{\varphi^{n_m}\}_{m=1}^{\infty}$ of $\{\varphi^n\}_{n=1}^{\infty}$ such that $\lim_{m\to\infty} \sigma(\varphi^{n_m}) < \sigma(\varphi)$. Let $t \in \mathbb{R}$ be such that $\lim_{m\to\infty} \sigma(\varphi^{n_m}) < t < \sigma(\varphi)$. For each $p \in \Delta$ such that $d(p, K) \leq t$, since $\sigma(\varphi) = \min_{q\in B(\varphi)} d(q, K) > t$, $p \notin B(\varphi)$, and thus, $E_p\varphi \neq I(\varphi)$. Since $f \sim x_{*f}$, by (22), $I(\varphi) = I(u(f)) = I(u(x_{*f})) = I(\varphi_*\mathbf{1})$, and since I is normalized, $I(\varphi) = \varphi_*$. For each $p \in \Delta$ such that $d(p, K) \leq t$, since $E_p\varphi \neq I(\varphi)$ and $E_p\varphi \geq \varphi_* = I(\varphi)$, $E_p\varphi > I(\varphi)$. Thus,

$$\min_{p\in\Delta:d(p,K)\leq t} E_p\varphi > I(\varphi). \tag{36}$$

Since $\lim_{m\to\infty} \sigma(\varphi^{n_m}) < t$, for sufficiently large m, $\min_{p\in\Delta:d(p,K)\leq\sigma(\varphi^{n_m})} E_p\varphi^{n_m} \geq \min_{p\in\Delta:d(p,K)\leq t} E_p\varphi^{n_m}$. Since I is continuous, and since for each $m \in \mathbb{N}$, $I(\varphi^{n_m}) = \min_{p\in\Delta:d(p,K)\leq\sigma(\varphi^{n_m})} E_p\varphi^{n_m}$, $I(\varphi) = \lim_{m\to\infty} I(\varphi^{n_m}) = \lim_{m\to\infty} \min_{p\in\Delta:d(p,K)\leq\sigma(\varphi^{n_m})} E_p\varphi^{n_m}$. Thus,

$$\begin{aligned} I(\varphi) &= \lim_{m\to\infty} \min_{p\in\Delta:d(p,K)\leq\sigma(\varphi^{n_m})} E_p\varphi^{n_m} \geq \lim_{m\to\infty} \min_{p\in\Delta:d(p,K)\leq t} E_p\varphi^{n_m} \\ &= \min_{p\in\Delta:d(p,K)\leq t} E_p\varphi, \end{aligned}$$

which contradicts (36), as desired.

We complete the proof of the sufficiency of the axioms by showing that K is actually the largest essential set, so that $\langle u, K, \sigma\rangle$ is a canonical VC representation of $\succsim$. Let $\langle u', K', \sigma'\rangle$ be another VC representation of $\succsim$. To show that $K' \subseteq K$, let $p' \in K'$ and $\varphi \in \mathbb{R}^S$. By the definition of K, it is equivalent for us to show that $I(\varphi) \leq E_{p'}\varphi$. Let $f \in \mathcal{F}$ such that $u(f) = \varphi$. Since $\langle u', K', \sigma'\rangle$ is a VC representation of $\succsim$, $u'(x_f) = \min_{p\in\Delta:d(p,K')\leq\sigma'(u'(f))} E_p u'(f)$. Since $p' \in K'$, $\min_{p\in\Delta:d(p,K')\leq\sigma'(u'(f))} E_p u'(f) \leq E_{p'}u'(f)$. Thus, $u'(x_f) \leq E_{p'}u'(f)$. Since both u and u' are affine functions representing $\succsim$ restricted to X, by routine arguments, u' is a positive affine transformation of u. Thus, $u(x_f) \leq E_{p'}u(f)$. Since I is normalized, by (22), $I(\varphi) = I(u(f)) = I(u(x_f)) = u(x_f)$. Hence, $I(\varphi) \leq E_{p'}\varphi$.

To prove the necessity of the axioms, let $\succsim$ admit a canonical VC representation $\langle u, K, \sigma\rangle$ with $\sigma \in \Sigma(K)$. Clearly, it satisfies A.1. Since u is affine, it satisfies A.2. Since $\sigma \in \Sigma(K)$, one can readily verify that it satisfies A.4. Since u is onto, by Lemma 29 of Maccheroni et al. (2006), it satisfies A.6.

To show that $\succsim$ satisfies A.3, define $C : \mathbb{R}^S \rightrightarrows \Delta$ and $J : \mathbb{R}^S \to \mathbb{R}$ by setting for each $\varphi \in \mathbb{R}^S$,

$$C(\varphi) := \{p \in \Delta : d(p, K) \leq \sigma(\varphi)\},$$

and

$$J(\varphi) := \min_{p \in C(\varphi)} E_p \varphi,$$

and it suffices to show that J is continuous. Let $\varphi \in \mathbb{R}^S$ and $\{\varphi^n\}_{n=1}^{\infty}$ be a sequence of elements in $\mathbb{R}^S$ such that $\lim_{n\to\infty} \varphi^n = \varphi$. If φ is constant, then for each $n \in \mathbb{N}$, $|J(\varphi^n) - J(\varphi)| \leq \sup_{s\in S} |\varphi^n(s) - \varphi(s)|$, and thus, $0 \leq \lim_{n\to\infty} |J(\varphi^n) - J(\varphi)| \leq \lim_{n\to\infty} \sup_{s\in S} |\varphi^n(s) - \varphi(s)| = 0$, so that $\lim_{n\to\infty} J(\varphi^n) = J(\varphi)$. Suppose that φ is not constant. Let $f \in \mathcal{F}$ be such that $u(f) = \varphi$. We shall show that $\lim_{n\to\infty} J(\varphi^n) = J(\varphi)$ in each of the following two cases.

Case 1: $f \nsim x_{*f}$. Then σ is continuous at φ. To show the continuity of J at φ, by the maximum theorem, it suffices to show the continuity of the correspondence C at φ.

To show the upper hemicontinuity of C at φ, let $\bar{p} \in \Delta$ and let $\{p^n\}_{n=1}^{\infty}$ be a sequence of elements in Δ such that $\lim_{n\to\infty} p^n = \bar{p}$ and for each $n \in \mathbb{N}$, $p^n \in C(\varphi^n)$. We shall show that $\bar{p} \in C(\varphi)$. For each $n \in \mathbb{N}$, since $p^n \in C(\varphi^n)$, $d(p^n, K) \leq \sigma(\varphi^n)$. Since $d(\cdot, K)$ is continuous on Δ and σ is continuous at φ, $d(\bar{p}, K) = \lim_{n\to\infty} d(p^n, K)$ and $\sigma(\varphi) = \lim_{n\to\infty} \sigma(\varphi^n)$. Thus, $d(\bar{p}, K) \leq \sigma(\varphi)$. Hence, $p \in C(\varphi)$.

To show the lower hemicontinuity of C at φ, let $\bar{p} \in C(\varphi)$. We shall show that there is a sequence $\{p^n\}_{n=1}^{\infty}$ such that $\lim_{n\to\infty} p^n = \bar{p}$ and for each $n \in \mathbb{N}$, $p^n \in C(\varphi^n)$. Since $\bar{p} \in C(\varphi)$, $d(\bar{p}, K) \leq \sigma(\varphi)$. Let $q \in K$ be such that $d(\bar{p}, q) = d(\bar{p}, K)$. Then $d(\bar{p}, q) \leq \sigma(\varphi)$. For each $n \in \mathbb{N}$, let

$$p^n := \begin{cases} \bar{p} & \text{if } d(\bar{p}, q) \leq \sigma(\varphi^n), \\ \epsilon^n q + (1 - \epsilon^n)\bar{p} & \text{if } d(\bar{p}, q) > \sigma(\varphi^n), \end{cases}$$

where $\epsilon^n \in (0, 1]$ is such that $d(\epsilon^n q + (1 - \epsilon^n)\bar{p}, q) = \sigma(\varphi^n)$. Then for each $n \in \mathbb{N}$, $d(p^n, K) \leq d(p^n, q) \leq \sigma(\varphi^n)$, and thus, $p^n \in C(\varphi^n)$. Moreover, for each $n \in \mathbb{N}$, if $p^n = \bar{p}$, then $d(\bar{p}, p^n) = 0$; if $p^n = \epsilon^n q + (1 - \epsilon^n)\bar{p}$, then $d(p^n, q) = d(\epsilon^n q + (1-\epsilon^n)\bar{p}, q) = \sigma(\varphi^n) < d(\bar{p}, q)$, and thus $d(\bar{p}, p^n) = d(\bar{p}, q) - d(p^n, q) = d(\bar{p}, q) - \sigma(\varphi^n) \leq \sigma(\varphi) - \sigma(\varphi^n)$. Then by the continuity of σ at φ, $\lim_{n\to\infty} p^n = \bar{p}$.

Case 2: $f \sim x_{*f}$. Then σ is lower semicontinuous at φ. To show the continuity of J at φ, suppose to the contrary that there exist $\epsilon > 0$ and a subsequence $\{\varphi^{n_m}\}_{m=1}^{\infty}$ of $\{\varphi^n\}_{n=1}^{\infty}$ such that for each $m \in \mathbb{N}$, $|J(\varphi^{n_m}) - J(\varphi)| > \epsilon$. Since $f \sim x_{*f}$, and since $\langle u, K, \sigma \rangle$ is a VC representation of $\succsim$, $J(u(f)) = J(u(x_{*f})) = \varphi_*$. Thus, $J(\varphi) = \varphi_*$, and for each $m \in \mathbb{N}$, $|J(\varphi^{n_m}) - \varphi_*| > \epsilon$. Then for sufficiently large m, $|J(\varphi^{n_m}) - \varphi_*^{n_m}| > \frac{\epsilon}{2}$, and since J is non-decreasing, $J(\varphi^{n_m}) - \varphi_*^{n_m} = |J(\varphi^{n_m}) - \varphi_*^{n_m}| > \frac{\epsilon}{2}$. Let $\alpha \in (0, 1)$ be such that $\alpha(\varphi^* - \varphi_*) < \frac{\epsilon}{2}$. Then for sufficiently large m, $\alpha(\varphi^{n_m *} - \varphi_*^{n_m}) < \frac{\epsilon}{2} < J(\varphi^{n_m}) - \varphi_*^{n_m}$, and thus, $(1-\alpha)\varphi_*^{n_m} + \alpha\varphi^{n_m *} < J(\varphi^{n_m})$. Let

$B^{\infty} := \{p \in \Delta : E_p\varphi = (1-\alpha)\varphi_* + \alpha\varphi^*\}$, and for each $m \in \mathbb{N}$, $B^m := \{p \in \Delta : E_p\varphi^{n_m} = (1-\alpha)\varphi_*^{n_m} + \alpha\varphi^{n_m *}\}$. Observe that B^{∞} is non-empty and closed, and so is B^m for each $m \in \mathbb{N}$. Moreover, for sufficiently large m, if $p \in B^m$, then $E_p\varphi^{n_m} = (1-\alpha)\varphi_*^{n_m} + \alpha\varphi^{n_m *} < J(\varphi^{n_m})$, so by the definition of J, $p \notin C(\varphi^{n_m})$, and thus, by the definition of C, $d(p, K) > \sigma(\varphi^{n_m})$. Hence, for sufficiently large m, $\min_{p \in B^m} d(p, K) > \sigma(\varphi^{n_m})$. Thus, $\liminf_{m\to\infty} \min_{p\in B^m} d(p, K) \geq \liminf_{m\to\infty} \sigma(\varphi^{n_m}) \geq \liminf_{n\to\infty} \sigma(\varphi^n)$. Since φ is not constant, $\min_{p\in B^{\infty}} d(p, K) = \lim_{m\to\infty} \min_{p\in B^m} d(p, K)$, using similar arguments as in showing the continuity of σ defined by (35) in the case of $\varphi_* < I(\varphi) < \varphi^*$.[25] Since σ is lower semicontinuous at φ, $\liminf_{n\to\infty} \sigma(\varphi^n) \geq \sigma(\varphi)$. Thus, $\min_{p\in B^{\infty}} d(p, K) \geq \sigma(\varphi)$. Let $c := \min_{p\in B^{\infty}} d(p, K)$. Then $c \geq \sigma(\varphi)$. Hence,

$$\min_{p\in\Delta:d(p,K)\leq c} E_p\varphi \leq \min_{p\in C(\varphi)} E_p\varphi = J(\varphi) = \varphi_* < (1-\alpha)\varphi_* + \alpha\varphi^*, \tag{37}$$

where the last inequality holds since φ is not constant and $\alpha \in (0, 1)$. Observe that for each $p \in K$ and each $m \in \mathbb{N}$, $p \in C(\varphi^{n_m})$, so that $J(\varphi^{n_m}) \leq E_p\varphi^{n_m}$. Thus, for each $p \in K$, when m is sufficiently large, $(1-\alpha)\varphi_*^{n_m} + \alpha\varphi^{n_m *} < J(\varphi^{n_m}) \leq E_p\varphi^{n_m}$, and hence $E_p\varphi = \lim_{m\to\infty} E_p\varphi^{n_m} \geq \lim_{m\to\infty}(1-\alpha)\varphi_*^{n_m} + \alpha\varphi^{n_m *} = (1-\alpha)\varphi_* + \alpha\varphi^*$. By lemma 2, $\min_{p\in\Delta:d(p,K)\leq c} E_p\varphi = (1-\alpha)\varphi_* + \alpha\varphi^*$, which contradicts (37), as desired.

To show that $\succsim$ satisfies A.5.2, let $q \in K$. Let $\succsim_q$ be the SEU preference relation defined by setting for each pair $f, g \in \mathcal{F}$, $f \succsim_q g \iff \sum_{s\in S} q_s f(s) \succsim \sum_{s\in S} q_s g(s)$. We shall show that $\succsim$ is more ambiguity averse than the SEU preference relation $\succsim_q$. Let $f \in \mathcal{F}$ and $x \in X$ be such that $x \succsim_q f$. Then $x \succsim \sum_{s\in S} q_s f(s)$. Since $\langle u, K, \sigma\rangle$ is a varaint constraint representation of $\succsim$ and since u is affine, $u(x) \geq u(\sum_{s\in S} q_s f(s)) = E_q u(f)$. Since $q \in K$, $E_q u(f) \geq \min_{p\in\Delta:d(p,K)\leq\sigma(u(f))} E_p u(f)$. Then $u(x) \geq \min_{p\in\Delta:d(p,K)\leq\sigma(u(f))} E_p u(f)$, and thus, $x \succsim f$.

Lastly, to prove the uniqueness of the representation, let $\langle u, K, \sigma\rangle$ and $\langle u', K', \sigma'\rangle$ be two canonical VC representations of some preference relation $\succsim$. By routine arguments, u' is a positive affine transformation of u. By the definition of a canonical VC representation, K is the largest essential set, and so is K'. Thus, $K = K'$. As argued in the proof of the sufficiency of the axioms, K coincides with the benchmark set. Finally, suppose that $u = u'$, let $f \in \mathcal{F}$ be such that $f \nsim x_{*f}$, and we want to show that $\sigma(u(f)) = \sigma'(u'(f))$. Let $\varphi := u(f)$. Suppose that $\sigma(\varphi) < \sigma'(\varphi)$. Let $p_*, q \in \Delta$ be such that $E_{p_*}\varphi = \varphi_*$, $E_q\varphi = \min_{p\in\Delta:d(p,K)\leq\sigma(\varphi)} E_p\varphi$, and $d(q, K) \leq \sigma(\varphi)$. Since $\langle u, K, \sigma\rangle$ is a VC representation of $\succsim$, and since $f \nsim x_{*f}$, $u(x_f) = \min_{p\in\Delta:d(p,K)\leq\sigma(u(f))} E_p u(f) \neq u(x_{*f}) = \varphi_*$. Thus, $E_q\varphi = u(x_f) \neq \varphi_*$. Since $d(q, K) \leq \sigma(\varphi) < \sigma'(\varphi)$ and $d(\cdot, K)$ is continuous on Δ, there

[25] In the proof of the sufficiency of the axioms, the arguments used in showing the continuity of σ at φ in the case $\varphi_* < I(\varphi) < \varphi^*$ (Case 2) can be applied here by taking for each $\varphi' \in \mathbb{R}^S$, $I(\varphi') = \alpha\varphi'_* + (1-\alpha)\varphi'^*$.

is $\epsilon \in (0,1)$ such that $d(\epsilon p_* + (1-\epsilon)q, K) < \sigma'(\varphi)$. Thus, $\min_{p \in \Delta: d(p,K) \leq \sigma'(\varphi)} E_p\varphi \leq E_{\epsilon p_* + (1-\epsilon)q}\varphi = \epsilon\varphi_* + (1-\epsilon)E_q\varphi \neq u(x_f)$. Since $\langle u', K', \sigma' \rangle$ is a VC representation of $\succsim$, and since $u = u'$, $K = K'$, and $\min_{p \in \Delta: d(p,K) \leq \sigma'(\varphi)} E_p\varphi \neq u(x_f)$, $f \nsim x_f$, which is a contradiction. By analogous arguments, $\sigma(\varphi) > \sigma'(\varphi)$ is not possible. Hence, $\sigma(\varphi) = \sigma'(\varphi)$, or equivalently, $\sigma(u(f)) = \sigma(u'(f))$. □

Proof of Proposition 2 Suppose that $\succsim$ admits both a canonical WM representation $\langle u, D, \lambda \rangle$ and a canonical VC representation $\langle u', K, \sigma \rangle$. To show that $K \subseteq D$, suppose to the contrary that there is $q \in K \backslash D$. Since $q \notin D$, by a standard separation theorem, there is $\varphi \in \mathbb{R}^S$ such that $E_q\varphi < \min_{p \in D} E_p\varphi$. Let $f \in \mathcal{F}$ and $x \in X$ be such that $u(f) = \varphi$ and $u(x) = \min_{p \in D} E_p\varphi$. Since both u and u' are affine functions representing $\succsim$ restricted to X, by routine arguments, u' is a positive affine transformation of u. Thus, $u'(x) = \min_{p \in D} E_p u'(f) > E_q u'(f)$. Since $q \in K$, $E_q u'(f) \geq \min_{p \in \Delta: d(p,K) \leq \sigma(u'(f))} E_p u'(f)$. Hence, $u'(x) > \min_{p \in \Delta: d(p,K) \leq \sigma(u'(f))} E_p u'(f)$, and since $\langle u', K, \sigma \rangle$ is a VC representation of $\succsim$, $x \succ f$. Since $\langle u, D, \lambda \rangle$ is a WM representation of $\succsim$, and since

$$
\begin{aligned}
u(x) = \min_{p \in D} E_p\varphi = \min_{p \in D} E_p u(f) \leq \lambda(u(f)) \min_{p \in D} E_p u(f) \\
+ (1 - \lambda(u(f))) \max_{p \in D} E_p u(f),
\end{aligned}
$$

$f \succsim x$, which contradicts $x \succ f$, as desired. □

Proof of Proposition 3 To prove the "only if" direction, let $\succsim$ be a MEU preference relation. Thus, there exist an affine onto function $u : \mathbb{R}^S \to \mathbb{R}$ and a non-empty, closed, and convex set $D \subseteq \Delta$ such that for each pair $f, g \in \mathcal{F}$, $f \succsim g \iff \min_{p \in D} E_p u(f) \geq \min_{p \in D} E_p u(g)$. It is known that a MEU preference relation satisfies A.1–A.4, A.5.1, A.5.2, and A.6. By Ghirardato et al. (2004), $\succsim$ admits a Bewley WM representation with D being the Bewley set.[26] By Theorem 1, it admits a canonical WM representation with the smallest admissible set being the Bewley set. Thus, D is the smallest admissible set, and when $\lambda : \mathbb{R}^S \to [0,1]$ is a constant function equal to 1, $\langle u, D, \lambda \rangle$ is a canonical WM representation of $\succsim$. Moreover, by Ghirardato and Marinacci (2002), D is the benchmark set.[27] By Theorem 2, $\succsim$ admits a canonical VC representation with the largest essential set being the benchmark set. Thus, D is the largest essential set, and when $\sigma : \mathbb{R}^S \to \mathbb{R}_+$ is a constant function equal to 0, $\langle u, D, \sigma \rangle$ is a canonical VC representation of $\succsim$.

To prove the "if" direction, let $\succsim$ admit both a canonical WM representation $\langle u, D, \lambda \rangle$ and a canonical VC representation $\langle u', K, \sigma \rangle$ with $K = D$. To show

[26] See their Theorem 11 and the first paragraph of their Section 5.1. The same result was reported by Klaus Nehring in his talk "Preference and Belief without the Independence Axiom" at the LOFT2 conference in Torino (Italy), December 1996.

[27] See their Corollary 14.

that $\succsim$ is a MEU preference relation, it suffices to show that for each $f \in \mathcal{F}$, $\lambda(u(f)) = 1$. Let $f \in \mathcal{F}$. Since $\langle u', K, \sigma\rangle$ is a VC representation of $\succsim$, $u'(x_f) = \min_{p\in\Delta:d(p,K)\leq\sigma(u'(f))} E_p u'(f)$. Since both u and u' are affine functions representing $\succsim$ restricted to X, by routine arguments, u is a positive affine transformation of u'. Thus, $u(x_f) = \min_{p\in\Delta:d(p,K)\leq\sigma(u'(f))} E_p u(f) \leq \min_{p\in K} E_p u(f)$. Since $\langle u, D, \lambda\rangle$ is a WM representation of $\succsim$, and since $K = D$, $u(x_f) = \lambda(u(f)) \min_{p\in K} E_p u(f) + (1 - \lambda(u(f))) \max_{p\in K} E_p u(f) \geq \min_{p\in K} E_p u(f)$. Thus, $u(x_f) = \min_{p\in K} E_p u(f)$, or equivalently, $\lambda(u(f)) = 1$. □

Proof of Corollary 2 Let $\succsim$ admit an ambiguity averse representation $\langle u, G\rangle$. Define $I : \mathbb{R}^S \to \mathbb{R}$ by setting for each $\varphi \in \mathbb{R}^S$, $I(\varphi) := \min_{p\in\Delta} G(E_p\varphi, p)$. By the proof of Theorem 3 of Cerreia-Vioglio et al. (2011b), I is normalized, non-decreasing, and continuous. Moreover, by their Proposition 11, $K^* = \bigcap_{x\in X} \pi(x)$ where $\pi(\cdot)$ is defined in (17). Thus, $p \in K^*$ if and only if for each $x \in X$ and each $f \in \mathcal{F}$, $u(x) \geq E_p u(f) \Longrightarrow x \succsim f$. Then by definition, K^* is the benchmark set. Since K^* is the benchmark set and $K^* \neq \emptyset$, $\succsim$ satisfies A.5.2. Thus, by the proof of Theorem 2, $\succsim$ admits a canonical VC representation $\langle u, K^*, \sigma\rangle$ in which σ belongs to $\Sigma(K^*)$ and is given by (35). □

Proof of Proposition 4 Let $\succsim$ admit a multiplier representation $\langle u, q, \theta\rangle$. Equivalently, $\succsim$ admits an ambiguity averse representation $\langle u, G\rangle$ such that for each $(t, p) \in \mathbb{R}\times\Delta$, $G(t, p) = t + \theta R(p||q)$.[28] It can be readily seen that $K^* = \{q\}$, and

$$D^* = \begin{cases} \{p \in \Delta : p \ll q\} & \text{if } \theta < \infty, \\ \{q\} & \text{if } \theta = \infty. \end{cases}$$

Define $I : \mathbb{R}^S \to \mathbb{R}$ by setting for each $\varphi \in \mathbb{R}^S$, $I(\varphi) := \min_{p\in\Delta}[E_p\varphi + \theta R(p||q)]$. When $\theta < \infty$, by the variational formula (see, e.g., Proposition 1.4.2 of Dupuis and Ellis (1997)), for each $\varphi \in \mathbb{R}^S$,

$$I(\varphi) = -\theta \log E_q e^{-\frac{\varphi}{\theta}}. \tag{38}$$

By the proof of Corollary 1, $\succsim$ admits a canonical WM representation $\langle u, D^*, \lambda\rangle$ in which λ is given by (24) on $\{\varphi \in \mathbb{R}^S : \min_{p\in D^*} E_p\varphi \neq \max_{p\in D^*} E_p\varphi\}$. Suppose that $\theta < \infty$. Then $D^* = \{p \in \Delta : p \ll q\}$ and for each $\varphi \in \mathbb{R}^S$,

$$\min_{p\in D^*} E_p\varphi = \min_{s\in S:q_s>0} \varphi(s) \text{ and } \max_{p\in D^*} E_p\varphi = \max_{s\in S:q_s>0} \varphi(s). \tag{39}$$

For each $\varphi \in \mathbb{R}^S$, if $\min_{p\in D^*} E_p\varphi \neq \max_{p\in D^*} E_p\varphi$, then by substituting (38) and (39) into (24), we know that the value of $\lambda(\varphi)$ is given by (10), and if $\min_{s\in S:q_s>0} \varphi(s) =$

[28] See Theorem 24 of Cerreia-Vioglio et al. (2011b).

$\max_{s\in S:q_s>0} \varphi(s)$, it can be readily seen that $\lambda(\varphi)$ can take any value including that given by (10). Suppose that $\theta = \infty$. Then $D^* = \{q\}$ and for each $\varphi \in \mathbb{R}^S$, $\min_{p\in D^*} E_p\varphi = \max_{p\in D^*} E_p\varphi = E_q\varphi = I(\varphi)$. It can be readily seen that λ can take any value and we simply let λ be a constant function equal to 1.

Since $K^* = \{q\} \neq \emptyset$, by the proof of Corollary 2, $\succsim$ admits a canonical VC representation $\langle u, K^*, \sigma\rangle$ in which σ is given by (35). For each $\varphi \in \mathbb{R}^S$, if $\theta < \infty$, then by substituting (38) into (35), we know that the value of $\sigma(\varphi)$ is given by (11), and if $\theta = \infty$, then by substituting $I(\varphi) = E_q\varphi$ into (35), $\sigma(\varphi) = 0$. □

Proofs in Section 4

Lemma 3 *A preference relation $\succsim$ satisfies A.1, A.2.1, A.3, A.4, A.5.1, and A.6 if and only if there exist an affine onto function $u : X \to \mathbb{R}$ and a normalized, non-decreasing, quasi-concave, continuous, and constant superadditive functional $I : \mathbb{R}^S \to \mathbb{R}$ satisfying (22). Moreover, u is unique up to a positive affine transformation, and given u, there is a unique normalized functional $I : \mathbb{R}^S \to \mathbb{R}$ satisfying (22).*

Proof To prove the sufficiency of the axioms, let $\succsim$ satisfy A.1, A.2.1, A.3, A.4, A.5.1, and A.6. Note that A.2.1 implies A.2. Then by applying the same techniques of Lemma 57 of Cerreia-Vioglio et al. (2011b), one can show that there exist a non-constant affine function $u : X \to \mathbb{R}$ and a normalized, non-decreasing, quasi-concave, and continuous functional $I : u(X)^S \to \mathbb{R}$ satisfying (22). Moreover, since $\succsim$ satisfies A.6, by Lemma 29 of Maccheroni et al. (2006), $u(X) = \mathbb{R}$. Thus, we only need to check that I is constant superadditive.

Let $\varphi \in \mathbb{R}^S$ and $t \in \mathbb{R}_+$. Since $u(X) = \mathbb{R}$, there are $x, y \in X$ and $f \in \mathcal{F}$ such that $u(x) = 0$, $u(y) = 2t$, and $u(f) = 2\varphi$. Then $u(\frac{1}{2}f + \frac{1}{2}y) = \varphi + t\mathbf{1}$ and $u(\frac{1}{2}f + \frac{1}{2}x) = \varphi$. Since u is affine and onto, there is $z \in X$ satisfying $u(\frac{1}{2}f + \frac{1}{2}x) = u(\frac{1}{2}z + \frac{1}{2}x)$, so that $\frac{1}{2}f + \frac{1}{2}x \sim \frac{1}{2}z + \frac{1}{2}x$. Since $t \geq 0$, $y \succsim x$. Then by A.2.1, $\frac{1}{2}f + \frac{1}{2}y \succsim \frac{1}{2}z + \frac{1}{2}y$. Thus,

$$\begin{aligned} I(\varphi + t\mathbf{1}) &= I\left(u\left(\frac{1}{2}f + \frac{1}{2}y\right)\right) \geq I\left(u\left(\frac{1}{2}z + \frac{1}{2}y\right)\right) \\ &= \frac{1}{2}u(z) + \frac{1}{2}u(y) = \frac{1}{2}u(z) + \frac{1}{2}u(x) + \frac{1}{2}u(y) = u\left(\frac{1}{2}z + \frac{1}{2}x\right) + t \\ &= I\left(u\left(\frac{1}{2}f + \frac{1}{2}x\right)\right) + t = I(\varphi) + t. \end{aligned}$$

Hence, I is constant superadditive.

The necessity of the axioms can be readily seen, and the uniqueness follows from routine arguments. Thus, we omit the proofs. □

Proof of Theorem 3 Assume first that $\succsim$ satisfies A.1, A.2.1, A.3, A.4, A.5.1, and A.6, i.e., statement 1 holds. We will prove three representation results, respectively, in statements 2, 3, and 4. When proving each representation result, we will also prove the additional properties of the representation stated in the end of Theorem 3.

Since $\succsim$ satisfies A.1, A.2.1, A.3, A.4, A.5.1, and A.6, by Lemma 3, there exist an affine onto function $u : X \to \mathbb{R}$ and a normalized, non-decreasing, quasi-concave, continuous, and constant superadditive functional $I : \mathbb{R}^S \to \mathbb{R}$ satisfying (22).

We first prove that statement 2 holds. The following observations will be useful. Let $\varphi \in \mathbb{R}^S$. First, for each $k > 0$ and each $t \in \mathbb{R}$, $\frac{1}{k}[I(k\varphi - t\mathbf{1}) + t] \in [\varphi_*, \varphi^*]$. This can be readily shown by the fact that I is normalized and non-decreasing. Second, for each $k > 0$, $\frac{1}{k}[I(k\varphi - t\mathbf{1}) + t]$ is non-increasing in t on $\mathbb{R}$ since I is constant superadditive. Third, $\lim_{t\to\infty} \frac{1}{k}[I(k\varphi - t\mathbf{1}) + t]$ is non-increasing in k on $(0, \infty)$. To see this, suppose to the contrary that there are $k, k' \in (0, \infty)$ such that $k < k'$ and $\lim_{t\to\infty} \frac{1}{k}[I(k\varphi - t\mathbf{1}) + t] < \lim_{t\to\infty} \frac{1}{k'}[I(k'\varphi - t\mathbf{1}) + t]$. Hence, there is $\bar{t} \in \mathbb{R}$ such that for each pair $t, t' \in [\bar{t}, \infty)$, $\frac{1}{k}[I(k\varphi - t\mathbf{1}) + t] < \frac{1}{k'}[I(k'\varphi - t'\mathbf{1}) + t']$. That is,

$$I(k\varphi - t\mathbf{1}) < \frac{k}{k'} I(k'\varphi - t'\mathbf{1}) - t + \frac{k}{k'} t'. \tag{40}$$

Pick $t, t' \geq \bar{t}$ such that $\frac{k}{k'}(-t') + (1 - \frac{k}{k'}) I(k'\varphi - t'\mathbf{1}) = -t$. Thus, $k\varphi - t\mathbf{1} = \frac{k}{k'}(k'\varphi - t'\mathbf{1}) + (1 - \frac{k}{k'}) I(k'\varphi - t'\mathbf{1})\mathbf{1}$. Since I is normalized and quasi-concave, and by the choice of t, t', we have

$$I(k\varphi - t\mathbf{1}) \geq I(k'\varphi - t'\mathbf{1}) = \frac{k}{k'} I(k'\varphi - t'\mathbf{1}) - t + \frac{k}{k'} t',$$

which contradicts (40), as desired.

Define $J : \mathbb{R}^S \to \mathbb{R}$ by setting for each $\varphi \in \mathbb{R}^S$,

$$J(\varphi) := \lim_{k\to\infty} \lim_{t\to\infty} \frac{1}{k}[I(k\varphi - t\mathbf{1}) + t]. \tag{41}$$

Because of the three observations above, J is well-defined, and for each $\varphi \in \mathbb{R}^S$, $I(\varphi) \geq J(\varphi)$. Since I is normalized, so is J. Since I is non-decreasing, so is J. By definition, J is constant additive and positive homogeneous of degree 1. We now check that J is superadditive. Suppose to the contrary that there are $\varphi, \varphi' \in \mathbb{R}^S$ such that $J(\varphi + \varphi') < J(\varphi) + J(\varphi')$. Since J is positive homogeneous of degree 1, $J(\frac{1}{2}\varphi + \frac{1}{2}\varphi') < \frac{1}{2}J(\varphi) + \frac{1}{2}J(\varphi')$. Thus, there is $k > 0$ and $\bar{t} \in \mathbb{R}$ such that whenever $t, t', t'' \geq \bar{t}$,

$$\frac{1}{k}\left[I\left(k\left(\frac{1}{2}\varphi + \frac{1}{2}\varphi'\right) - t''\mathbf{1}\right) + t''\right] < \frac{1}{2k}[I(k\varphi - t\mathbf{1}) + t] + \frac{1}{2k}[I(k\varphi' - t'\mathbf{1}) + t']. \tag{42}$$

Let $t, t' \geq \bar{t}$ be such that $I(k\varphi - t\mathbf{1}) = I(k\varphi' - t'\mathbf{1})$. Let $t'' = \frac{t+t'}{2}$. Then (42) becomes

$$I\left(\frac{1}{2}(k\varphi - t\mathbf{1}) + \frac{1}{2}(k\varphi' - t'\mathbf{1})\right) < \frac{1}{2}I(k\varphi - t\mathbf{1}) + \frac{1}{2}I(k\varphi' - t'\mathbf{1}). \tag{43}$$

Since $I(k\varphi - t\mathbf{1}) = I(k\varphi' - t'\mathbf{1})$ and I is quasi-concave,

$$I\left(\frac{1}{2}(k\varphi - t\mathbf{1}) + \frac{1}{2}(k\varphi' - t'\mathbf{1})\right) \geq \frac{1}{2}I(k\varphi - t\mathbf{1}) + \frac{1}{2}I(k\varphi' - t'\mathbf{1}),$$

which contradicts (43), as desired.

Since J is normalized, non-decreasing, constant additive, positive homogeneous of degree 1, and superadditive, by Lemma 3.5 of Gilboa and Schmeidler (1989), there is a unique non-empty, closed, and convex set $D \subseteq \Delta$ such that for each $\varphi \in \mathbb{R}^S$, $J(\varphi) = \min_{p \in D} E_p\varphi$.

Fix $\varphi \in \mathbb{R}^S$. We claim that $I(\varphi) \in [\min_{p \in D} E_p\varphi, \max_{p \in D} E_p\varphi]$. Since $I(\varphi) \geq J(\varphi)$ and $J(\varphi) = \min_{p \in D} E_p\varphi$, $I(\varphi) \geq \min_{p \in D} E_p\varphi$. To see that $I(\varphi) \leq \max_{p \in D} E_p\varphi$, let $t \in \mathbb{R}$ be such that $I(\varphi) = I(-\varphi + t\mathbf{1})$. Since $I(\varphi) = I(-\varphi + t\mathbf{1})$ and I is quasi-concave,

$$I\left(\frac{1}{2}\varphi + \frac{1}{2}(-\varphi + t\mathbf{1})\right) \geq \frac{1}{2}I(\varphi) + \frac{1}{2}I(-\varphi + t\mathbf{1}). \tag{44}$$

Since I is normalized, $I(\frac{1}{2}\varphi + \frac{1}{2}(-\varphi + t\mathbf{1})) = \frac{1}{2}t$. Since $I(-\varphi + t\mathbf{1}) \geq \min_{p \in D} E_p(-\varphi + t\mathbf{1})$, $\frac{1}{2}I(\varphi) + \frac{1}{2}I(-\varphi + t\mathbf{1}) \geq \frac{1}{2}I(\varphi) + \frac{1}{2}\min_{p \in D} E_p(-\varphi) + \frac{1}{2}t$. Thus, by (44),

$$\frac{1}{2}t \geq \frac{1}{2}I(\varphi) + \frac{1}{2}\min_{p \in D} E_p(-\varphi) + \frac{1}{2}t,$$

which implies that $I(\varphi) \leq -\min_{p \in D} E_p(-\varphi) = \max_{p \in D} E_p\varphi$.

Define $\lambda : \mathbb{R}^S \to [0, 1]$ by setting for each $\varphi \in \mathbb{R}^S$,

$$\lambda(\varphi) := \begin{cases} \dfrac{\max_{p \in D} E_p\varphi - I(\varphi)}{\max_{p \in D} E_p\varphi - \min_{p \in D} E_p\varphi} & \text{if } \min_{p \in D} E_p\varphi \neq \max_{p \in D} E_p\varphi, \\ 1 & \text{if } \min_{p \in D} E_p\varphi = \max_{p \in D} E_p\varphi. \end{cases}$$

By the definition of λ, for each $\varphi \in \mathbb{R}^S$, $I(\varphi) = \lambda(\varphi)\min_{p \in D} E_p\varphi + (1 - \lambda(\varphi))\max_{p \in D} E_p\varphi$. Since I is continuous, λ is continuous on $\{\varphi \in \mathbb{R}^S : \min_{p \in D} E_p\varphi \neq \max_{p \in D} E_p\varphi\}$. Thus, $\langle u, D, \lambda\rangle$ is a WM representation of $\succsim$. By the same arguments as in the first paragraph of the proof of Theorem 1, $\lambda \in \Lambda(D)$. Since I is constant superadditive, for each $\varphi \in \mathbb{R}^S$, $\lambda(\varphi + t\mathbf{1})$ is non-increasing in t on $\mathbb{R}$. By the third observation in the second paragraph, for each $\varphi \in \mathbb{R}^S$, $\lim_{t\to\infty} \lambda(k\varphi - t\mathbf{1})$ is non-decreasing in k on $(0, \infty)$. Moreover, since for each $\varphi \in \mathbb{R}^S$, $\lim_{k\to\infty}\lim_{t\to\infty} \frac{1}{k}[I(k\varphi - t\mathbf{1}) + t] = J(\varphi) = \min_{p \in D} E_p\varphi$, it can be readily shown that $\lim_{k\to\infty}\lim_{t\to\infty} \lambda(k\varphi - t\mathbf{1}) = 1$.

Now we show that D is the smallest admissible set. Since A.2.1 implies A.2, $\succsim$ satisfies A.1–A.4, A.5.1, and A.6. Thus, by Theorem 1, there is a canonical WM representation $\langle u', D', \lambda' \rangle$ of $\succsim$, where D' coincides with the Bewley set, and moreover, the uniqueness property of the representation implies that we can assume $u' = u$. To show that D is the smallest admissible set, it suffices to show that $D \subseteq D'$. Since D and D' are closed and convex, by Proposition A.1 of Ghirardato et al. (2004), it is sufficient to show that for each pair $\varphi, \varphi' \in \mathbb{R}^S$, if for each $p \in D'$, $E_p\varphi \geq E_p\varphi'$, then for each $p \in D$, $E_p\varphi \geq E_p\varphi'$.

Let $\varphi, \varphi' \in \mathbb{R}^S$ and $g, g' \in \mathcal{F}$ be such that $\varphi = u(g)$, $\varphi' = u(g')$, and for each $p \in D'$, $E_p\varphi \geq E_p\varphi'$. Let $\alpha \in (0, 1]$ and $h \in \mathcal{F}$. For each $k > 0$ and $t \in \mathbb{R}$, let $g_{k,t}, g'_{k,t} \in \mathcal{F}$ be such that $u(g_{k,t}) = ku(\alpha g + (1-\alpha)h) - t\mathbf{1}$ and $u(g'_{k,t}) = ku(\alpha g' + (1-\alpha)h) - t\mathbf{1}$. For each $k > 0$, each $t \in \mathbb{R}$, and each $p \in D'$, since $E_p u(g) = E_p\varphi \geq E_p\varphi' = E_p u(g')$, $E_p u(g_{k,t}) \geq E_p u(g'_{k,t})$. Since $\langle u, D', \lambda' \rangle$ is a WM representation of $\succsim$ with D' being the Bewley set, $\langle u, D' \rangle$ is a Bewley representation of the unambiguous preference relation $\succsim^*$ induced from $\succsim$. Thus, for each $k > 0$ and each $t \in \mathbb{R}$, $g_{k,t} \succsim^* g'_{k,t}$, and hence, $g_{k,t} \succsim g'_{k,t}$. For each $k > 0$ and each $t \in \mathbb{R}$, since $g_{k,t} \succsim g'_{k,t}$, $I(u(g_{k,t})) \geq I(u(g'_{k,t}))$, and thus, $I(ku(\alpha g + (1-\alpha)h) - t\mathbf{1}) \geq I(ku(\alpha g' + (1-\alpha)h) - t\mathbf{1})$. Hence,

$$\begin{aligned}
&J(u(\alpha g + (1-\alpha)h)) \\
&\quad = \lim_{k\to\infty} \lim_{t\to\infty} \frac{1}{k}[I(ku(\alpha g + (1-\alpha)h) - t\mathbf{1}) + t] \\
&\quad \geq \lim_{k\to\infty} \lim_{t\to\infty} \frac{1}{k}[I(ku(\alpha g' + (1-\alpha)h) - t\mathbf{1}) + t] \\
&\quad = J(u(\alpha g' + (1-\alpha)h)).
\end{aligned}$$

Define $\succsim'$ on $\mathcal{F}$ by setting for each pair $f, f' \in \mathcal{F}$,

$$f \succsim' f' \iff \min_{p\in D} E_p u(f) \geq \min_{p\in D} E_p u(f').$$

Let $\succsim'^*$ be the unambiguous preference relation induced from $\succsim'$. Since $J(u(\alpha g + (1-\alpha)h)) \geq J(u(\alpha g' + (1-\alpha)h))$, and since for each $f \in \mathcal{F}$, we have shown that $J(u(f)) = \min_{p\in D} E_p u(p)$, $\alpha g + (1-\alpha)h \succsim' \alpha g' + (1-\alpha)h$. Since $\alpha g + (1-\alpha)h \succsim' \alpha g' + (1-\alpha)h$, and since $\alpha \in (0, 1]$ and $h \in \mathcal{F}$ are arbitrarily chosen, $g \succsim'^* g'$. By Ghirardato et al. (2004),[29] for each $p \in D$, $E_p u(g) \geq E_p u(g')$, and thus, $E_p\varphi \geq E_p\varphi'$, as desired.

Next, we prove that statement 3 holds. By the similar arguments as in the second paragraph, we obtain the following observations. Let $\varphi \in \mathbb{R}^S$. First, for each $k > 0$ and each $t \in \mathbb{R}$, $\frac{1}{k}[I(k\varphi + t\mathbf{1}) - t] \in [\varphi_*, \varphi^*]$. Second, for each $k > 0$, $\frac{1}{k}[I(k\varphi + t\mathbf{1}) - t]$ is non-decreasing in t on $\mathbb{R}$. Third, $\lim_{t\to\infty} \frac{1}{k}[I(k\varphi + t\mathbf{1}) - t]$ is non-increasing in k on $(0, \infty)$.

[29] See the first paragraph on p.151 of Ghirardato et al. (2004).

Define $J' : \mathbb{R}^S \to \mathbb{R}$ by setting for each $\varphi \in \mathbb{R}^S$,

$$J'(\varphi) := \lim_{k \searrow 0} \lim_{t \to \infty} \frac{1}{k}[I(k\varphi + t\mathbf{1}) - t]. \tag{45}$$

Because of the three observations above, J' is well-defined, and for each $\varphi \in \mathbb{R}^S$, $I(\varphi) \leq J'(\varphi)$. Moreover, by the similar arguments as in the third paragraph, J' is normalized, non-decreasing, constant additive, positive homogeneous of degree 1, and superadditive. Thus, by Lemma 3.5 of Gilboa and Schmeidler (1989), there is a unique non-empty, closed, and convex set $K \subseteq \Delta$ such that for each $\varphi \in \mathbb{R}^S$, $J'(\varphi) = \min_{p \in K} E_p\varphi$.

We show that $K = \{p \in \Delta : \text{for each } \varphi \in \mathbb{R}^S, I(\varphi) \leq E_p\varphi\}$. Since for each $p' \in K$ and each $\varphi \in \mathbb{R}^S$, $I(\varphi) \leq J'(\varphi) = \min_{p \in K} E_p\varphi \leq E_{p'}\varphi$, $K \subseteq \{p \in \Delta :$ for each $\varphi \in \mathbb{R}^S, I(\varphi) \leq E_p\varphi\}$. To show that $\{p \in \Delta :$ for each $\varphi \in \mathbb{R}^S, I(\varphi) \leq E_p\varphi\} \subseteq K$, suppose to the contrary that there is $p' \in \Delta \backslash K$ such that for each $\varphi \in \mathbb{R}^S$, $I(\varphi) \leq E_{p'}\varphi$. Then for each $\varphi \in \mathbb{R}^S$, each $k > 0$, and each $t \in \mathbb{R}$, $I(k\varphi + t\mathbf{1}) \leq E_{p'}(k\varphi + t\mathbf{1}) = kE_{p'}\varphi + t$, and thus, $\frac{1}{k}[I(k\varphi + t\mathbf{1}) - t] \leq E_{p'}\varphi$. Hence, for each $\varphi \in \mathbb{R}^S$, $\min_{p \in K} E_p\varphi = J'(\varphi) = \lim_{k \searrow 0} \lim_{t \to \infty} \frac{1}{k}[I(k\varphi + t\mathbf{1}) - t] \leq E_{p'}\varphi$. However, since $p' \notin K$, by a standard separation theorem, there is $\varphi' \in \mathbb{R}^S$ such that $E_{p'}\varphi' < \min_{p \in K} E_p\varphi'$, which contradicts that for each $\varphi \in \mathbb{R}^S$, $\min_{p \in K} E_p\varphi \leq E_{p'}\varphi$, as desired.

Define $B : \mathbb{R}^S \rightrightarrows \Delta$ and $\sigma : \mathbb{R}^S \to \mathbb{R}_+$ by setting for each $\varphi \in \mathbb{R}^S$,

$$B(\varphi) := \{p \in \Delta : I(\varphi) = E_p\varphi\},$$

and

$$\sigma(\varphi) := \min_{p \in B(\varphi)} d(p, K).$$

Since K is non-empty and $K = \{p \in \Delta :$ for each $\varphi \in \mathbb{R}^S, I(\varphi) \leq E_p\varphi\}$, the latter set is also non-empty. Then by the same arguments as in the proof of the sufficiency of the axioms in Theorem 2, $\langle u, K, \sigma \rangle$ is a canonical VC representation of $\succsim$, and for each $\varphi \in \mathbb{R}^S$, $I(\varphi) = \min_{p \in \Delta : d(p,K) \leq \sigma(\varphi)} E_p\varphi$. Moreover, by the analogous arguments as in the first paragraph of the proof of Theorem 1, $\sigma \in \bar{\Sigma}(K)$.

To complete the proof that statement 1 implies statement 3, we fix an arbitrary $\varphi \in \mathbb{R}^S$ and show the following three properties of σ. First, $\sigma(\varphi + t\mathbf{1})$ is non-increasing in t on $\mathbb{R}$. Let $t, t' \in \mathbb{R}$ be such that $t \leq t'$. We want to show that $\sigma(\varphi + t\mathbf{1}) \geq \sigma(\varphi + t'\mathbf{1})$. Let $p^* \in B(\varphi + t\mathbf{1})$ be such that $d(p^*, K) = \min_{p \in B(\varphi + t\mathbf{1})} d(p, K)$. By the definition of σ, $d(p^*, K) = \sigma(\varphi + t\mathbf{1})$. Since $p^* \in B(\varphi + t\mathbf{1})$, by the definition of B, $I(\varphi + t\mathbf{1}) = E_{p^*}(\varphi + t\mathbf{1})$. Since $I(\varphi + t\mathbf{1}) = E_{p^*}(\varphi + t\mathbf{1})$ and $t \leq t'$, and since I is constant superadditive, $I(\varphi + t'\mathbf{1}) \geq E_{p^*}(\varphi + t'\mathbf{1})$. Assume first that $I(\varphi + t'\mathbf{1}) = E_{p^*}(\varphi + t'\mathbf{1})$. Then $p^* \in B(\varphi + t'\mathbf{1})$, and thus, $\sigma(\varphi + t\mathbf{1}) = d(p^*, K) \geq \min_{p \in B(\varphi + t'\mathbf{1})} d(p, K) = \sigma(\varphi + t'\mathbf{1})$, as desired. Assume now that $I(\varphi + t'\mathbf{1}) > E_{p^*}(\varphi + t'\mathbf{1})$. Suppose to the contrary that $\sigma(\varphi + t\mathbf{1}) < \sigma(\varphi + t'\mathbf{1})$. Then $d(p^*, K) = \sigma(\varphi + t\mathbf{1}) < \sigma(\varphi + t'\mathbf{1})$. Thus, $I(\varphi + t'\mathbf{1}) = \min_{p\in\Delta: d(p,K)\leq\sigma(\varphi+t'\mathbf{1})} E_p(\varphi + t'\mathbf{1}) \leq E_{p^*}(\varphi + t'\mathbf{1}) < I(\varphi + t'\mathbf{1})$, which is not possible. Hence, $\sigma(\varphi + t\mathbf{1}) \geq \sigma(\varphi + t'\mathbf{1})$.

Second, $\lim_{t\to\infty} \sigma(k\varphi + t\mathbf{1})$ is non-decreasing in k on $(0, \infty)$. For each $k > 0$, define $B_k : [0, \infty] \rightrightarrows \Delta$ by setting for each $t \in [0, \infty]$,

$$B_k(t) := \begin{cases} B(k\varphi + t\mathbf{1}) & \text{if } t < \infty, \\ \{p \in \Delta : \lim_{t\to\infty} \frac{1}{k}[I(k\varphi + t\mathbf{1}) - t] = E_p\varphi\} & \text{if } t = \infty. \end{cases}$$

Thus, for each $k > 0$ and each $t \in \mathbb{R}$, $\sigma(k\varphi + t\mathbf{1}) = \min_{p\in B_k(t)} d(p, K)$. Before proving the second property of σ, we first fix $k > 0$ and show that $\lim_{t\to\infty} \sigma(k\varphi + t\mathbf{1}) = \min_{p\in B_k(\infty)} d(p, K)$, or equivalently, $\lim_{t\to\infty} \min_{p\in B_k(t)} d(p, K) = \min_{p\in B_k(\infty)} d(p, K)$.

By the maximum theorem, it suffices to prove that B_k is continuous at ∞. The upper hemicontinuity can be readily seen, and thus, we omit the proof. For the lower hemicontinuity, let $\{t^n\}_{n=1}^\infty$ be a sequence of elements in $\mathbb{R}_+$ such that $\lim_{n\to\infty} t^n = \infty$. Fix $p^\infty \in B_k(\infty)$. We will show that there is a sequence $\{p^n\}_{n=1}^\infty$ of elements in Δ such that for each $n \in \mathbb{N}$, $p^n \in B_k(t^n)$, and $\lim_{n\to\infty} p^n = p^\infty$. Since $p^\infty \in B_k(\infty)$, $\lim_{t\to\infty} \frac{1}{k}[I(k\varphi+t\mathbf{1})-t] = E_{p^\infty}\varphi$. For each $n \in \mathbb{N}$, by the first observation in the first paragraph of our proof for statement 1 implying statement 3, $\frac{1}{k}[I(k\varphi+t^n\mathbf{1})-t^n] \geq \varphi_*$, and by the second observation, $\frac{1}{k}[I(k\varphi + t^n\mathbf{1}) - t^n] \leq \lim_{t\to\infty} \frac{1}{k}[I(k\varphi + t\mathbf{1}) - t] = E_{p^\infty}\varphi$. Let $p_* \in \Delta$ be such that $E_{p_*}\varphi = \varphi_*$. For each $n \in \mathbb{N}$, since $E_{p_*}\varphi = \varphi_* \leq \frac{1}{k}[I(k\varphi + t^n\mathbf{1}) - t^n] \leq E_{p^\infty}\varphi$, if $\varphi_* < E_{p^\infty}\varphi$, then there is a unique $\alpha^n \in [0, 1]$ satisfying that

$$\frac{1}{k}[I(k\varphi + t^n\mathbf{1}) - t^n] = E_{\alpha^n p_* + (1-\alpha^n)p^\infty}\varphi, \tag{46}$$

and if $\varphi_* = E_{p^\infty}\varphi$, each $\alpha^n \in [0, 1]$ satisfies (46), and we choose $\alpha^n := 0$ in this case. Let for each $n \in \mathbb{N}$, $p^n := \alpha^n p_* + (1 - \alpha^n)p^\infty$. Thus, for each $n \in \mathbb{N}$, $p^n \in \Delta$ and $\frac{1}{k}[I(k\varphi+t^n\mathbf{1})-t^n] = E_{p^n}\varphi$, the latter of which implies that $I(k\varphi+t^n\mathbf{1}) = E_{p^n}(k\varphi+t^n\mathbf{1})$, i.e., $p^n \in B_k(t_n)$. When $\varphi_* < E_{p^\infty}\varphi$, since $\lim_{t\to\infty} \frac{1}{k}[I(k\varphi+t\mathbf{1})-t] = E_{p^\infty}\varphi$, $\lim_{n\to\infty} p^n = p^\infty$. When $\varphi_* = E_{p^\infty}\varphi$, by definition, for each $n \in \mathbb{N}$, $p^n = p^\infty$, so that $\lim_{n\to\infty} p^n = p^\infty$.

To show that second property of σ, let $k, k' > 0$ be such that $k \leq k'$. We want to show that $\lim_{t\to\infty} \sigma(k\varphi + t\mathbf{1}) \leq \lim_{t\to\infty} \sigma(k'\varphi + t\mathbf{1})$. By the above result, it is equivalent to show that $\min_{p\in B_k(\infty)} d(p, K) \leq \min_{p\in B_{k'}(\infty)} d(p, K)$. Let $p^* \in B_k(\infty)$, $p^{*\prime} \in B_{k'}(\infty)$ and $q^*, q^{*\prime} \in K$ be such that $d(p^*, q^*) = \min_{p\in B_k(\infty)} d(p, K)$ and $d(p^{*\prime}, q^{*\prime}) = \min_{p\in B_{k'}(\infty)} d(p, K)$. We want to show that $d(p^*, q^*) \leq d(p^{*\prime}, q^{*\prime})$. Since $p^* \in B_k(\infty)$ and $p^{*\prime} \in B_{k'}(\infty)$, $\lim_{t\to\infty} \frac{1}{k}[I(k\varphi + t\mathbf{1}) - t] = E_{p^*}\varphi$ and $\lim_{t\to\infty} \frac{1}{k'}[I(k'\varphi + t\mathbf{1}) - t] = E_{p^{*\prime}}\varphi$. By the third observation, $\lim_{t\to\infty} \frac{1}{k'}[I(k'\varphi + t\mathbf{1})-t] \leq \lim_{t\to\infty} \frac{1}{k}[I(k\varphi+t\mathbf{1})-t]$, and thus, $E_{p^{*\prime}}\varphi \leq E_{p^*}\varphi$. By the definition of J' in (45) and the third observation again, $\lim_{t\to\infty} \frac{1}{k}[I(k\varphi + t\mathbf{1}) - t] \leq J'(\varphi)$, and thus,

$E_{p^*}\varphi \leq J'(\varphi)$. Since $J'(\varphi) = \min_{p\in K} E_p\varphi$ and $q^{*\prime} \in K$, $J'(\varphi) \leq E_{q^{*\prime}}\varphi$. Then $E_{p^{*\prime}}\varphi \leq E_{p^*}\varphi \leq J'(\varphi) \leq E_{q^{*\prime}}\varphi$. Hence, there is $\alpha \in [0, 1]$ such that $E_{p^*}\varphi = E_{\alpha p^{*\prime}+(1-\alpha)q^{*\prime}}\varphi$, and thus, $\lim_{t\to\infty} \frac{1}{k}[I(k\varphi + t\mathbf{1}) - t] = E_{\alpha p^{*\prime}+(1-\alpha)q^{*\prime}}\varphi$, i.e., $\alpha p^{*\prime} + (1-\alpha)q^{*\prime} \in B_k(\infty)$. Since $\alpha p^{*\prime}+(1-\alpha)q^{*\prime} \in B_k(\infty)$ and $q^{*\prime} \in K$, $\min_{p\in B_k(\infty)} d(p, K) \leq d(\alpha p^{*\prime}+(1-\alpha)q^{*\prime}, q^{*\prime})$. Thus, $d(p^*, q^*) = \min_{p\in B_k(\infty)} d(p, K) \leq d(\alpha p^{*\prime} + (1-\alpha)q^{*\prime}, q^{*\prime}) = \alpha d(p^{*\prime}, q^{*\prime}) \leq d(p^{*\prime}, q^{*\prime})$, as desired.

The last property of σ is that $\lim_{k\searrow 0} \lim_{t\to\infty} \sigma(k\varphi + t\mathbf{1}) = 0$. Let $p_* \in \Delta$ be such that $E_{p_*}\varphi = \varphi_*$. Let $q^* \in K$ be such that $E_{q^*}\varphi = \min_{p\in K} E_p\varphi$. For each $k > 0$ and each $t \in \mathbb{R}$, since $J'(k\varphi + t\mathbf{1}) = \min_{p\in K} E_p(k\varphi + t\mathbf{1}) = kE_{q^*}\varphi + t$ and $J'(k\varphi + t\mathbf{1}) \geq I(k\varphi + t\mathbf{1})$, $kE_{q^*}\varphi + t \geq I(k\varphi + t\mathbf{1})$; since I is normalized and non-decreasing, $I(k\varphi + t\mathbf{1}) \geq k\varphi_* + t$. Hence, for each $k > 0$ and each $t \in \mathbb{R}$, $kE_{q^*}\varphi + t \geq I(k\varphi + t\mathbf{1}) \geq k\varphi_* + t = kE_{p_*}\varphi + t$.

Assume that $E_{q^*}\varphi = E_{p_*}\varphi$. Then for each $k > 0$ and each $t \in \mathbb{R}$, $kE_{q^*}\varphi + t = kE_{p_*}\varphi + t$, and since $kE_{q^*}\varphi + t \geq I(k\varphi + t\mathbf{1}) \geq kE_{p_*}\varphi + t$, $I(k\varphi + t\mathbf{1}) = kE_{q^*}\varphi + t = E_{q^*}(k\varphi + t\mathbf{1})$, i.e., $q^* \in B(k\varphi + t\mathbf{1})$. Since for each $k > 0$ and each $t \in \mathbb{R}$, $q^* \in B(k\varphi + t\mathbf{1})$ and $q \in K$, $\sigma(k\varphi + t\mathbf{1}) = \min_{p\in B(k\varphi+t\mathbf{1})} d(p, K) = 0$. Thus, $\lim_{k\searrow 0} \lim_{t\to\infty} \sigma(k\varphi + t\mathbf{1}) = 0$.

Assume that $E_{q^*}\varphi > E_{p_*}\varphi$. For each $k > 0$ and each $t \in \mathbb{R}$, since $kE_{q^*}\varphi + t \geq I(k\varphi + t\mathbf{1}) \geq kE_{p_*}\varphi + t$, $E_{q^*}\varphi \geq \frac{1}{k}[I(k\varphi + t\mathbf{1}) - t] \geq E_{p_*}\varphi$. Hence, for each $k > 0$, $E_{q^*}\varphi \geq \lim_{t\to\infty} \frac{1}{k}[I(k\varphi + t\mathbf{1}) - t] \geq E_{p_*}\varphi$. Since $E_{q^*}\varphi > E_{p_*}\varphi$, for each $k > 0$,

$$\alpha_k := \frac{1}{E_{q^*}\varphi - E_{p_*}\varphi}\left[E_{q^*}\varphi - \lim_{t\to\infty} \frac{1}{k}[I(k\varphi + t\mathbf{1}) - t]\right]$$

is well-defined. It can be readily seen that for each $k > 0$, $\alpha_k \in [0, 1]$ and $E_{\alpha_k p_*+(1-\alpha_k)q^*}\varphi = \lim_{t\to\infty} \frac{1}{k}[I(k\varphi + t\mathbf{1}) - t]$. For each $k > 0$, recall the definition of $B_k(\infty)$ in the proof of the second property of σ, and recall our previous result that $\lim_{t\to\infty} \sigma(k\varphi + t\mathbf{1}) = \min_{p\in B_k(\infty)} d(p, K)$. For each $k > 0$, since $E_{\alpha_k p_*+(1-\alpha_k)q^*}\varphi = \lim_{t\to\infty} \frac{1}{k}[I(k\varphi + t\mathbf{1}) - t]$, $\alpha_k p_* + (1-\alpha_k)q^* \in B_k(\infty)$, and moreover, since $q^* \in K$, $\min_{p\in B_k(\infty)} d(p, K) \leq d(\alpha_k p_* + (1-\alpha_k)q^*, q^*)$. Thus, for each $k > 0$,

$$0 \leq \lim_{t\to\infty} \sigma(k\varphi + t\mathbf{1}) = \min_{p\in B_k(\infty)} d(p, K) \leq d(\alpha_k p_* + (1-\alpha_k)q^*, q^*) = \alpha_k d(p_*, q^*). \tag{47}$$

Since $J'(\varphi) = \min_{p\in K} E_p\varphi = E_{q^*}\varphi$ and by the definition of J' in (45), $\lim_{k\searrow 0} \lim_{t\to\infty} \frac{1}{k}[I(k\varphi + t\mathbf{1}) - t] = E_{q^*}\varphi$. Thus,

$$\lim_{k\searrow 0} \alpha_k = \frac{1}{E_{q^*}\varphi - E_{p_*}\varphi}\left[E_{q^*}\varphi - \lim_{k\searrow 0} \lim_{t\to\infty} \frac{1}{k}[I(k\varphi + t\mathbf{1}) - t]\right]$$
$$= \frac{1}{E_{q^*}\varphi - E_{p_*}\varphi}[E_{q^*}\varphi - E_{q^*}\varphi] = 0.$$

Thus, by taking the limits with respect to k on both sides of (47),

$$0 \leq \lim_{k \searrow 0} \lim_{t \to \infty} \sigma(k\varphi + t\mathbf{1}) \leq 0,$$

which implies that $\lim_{k \searrow 0} \lim_{t \to \infty} \sigma(k\varphi + t\mathbf{1}) = 0$, as desired.

Lastly, we prove that statement 4 holds. Since A.2.1 implies A.2, $\succsim$ satisfies A.1–A.4, A.5.1, and A.6. Then by Theorems 3 and 5 of Cerreia-Vioglio et al. (2011b), $\succsim$ admits an ambiguity averse representation $\langle u', G \rangle$, and the uniqueness property of their representation implies that we can assume $u' = u$. Moreover, they show that $G : \mathbb{R} \times \Delta \to (-\infty, \infty]$ is given by, for each $(t, p) \in \mathbb{R} \times \Delta$,

$$G(t, p) = \sup\{u(x_f) : f \in \mathcal{F}, E_p u(f) \leq t\}. \tag{48}$$

Fix $p \in \Delta$. Let $t, t' \in \mathbb{R}$ be such that $t \leq t'$. We want to show that $G(t, p) - t \leq G(t', p) - t'$. Let $f \in \mathcal{F}$ be such that $E_p u(f) \leq t$. Since u is onto, there is $f' \in \mathcal{F}$ satisfying that $u(f') = u(f) + (t' - t)\mathbf{1}$, so that $E_p u(f') \leq t'$. Since u and I satisfy (22), and since I is normalized and constant superadditive, $u(x_{f'}) = I(u(f')) = I(u(f) + (t' - t)\mathbf{1}) \geq I(u(f)) + t' - t = u(x_f) + t' - t$. Thus, $u(x_{f'}) - t' \geq u(x_f) - t$. Since for each $f \in \mathcal{F}$ such that $E_p u(f) \leq t$, there is $f' \in \mathcal{F}$ such that $E_p u(f') \leq t'$ and $u(x_{f'}) - t' \geq u(x_f) - t$, by (48), $G(t', p) - t' \geq G(t, p) - t$, as desired.

We now prove that statements 2, 3, and 4 all imply statement 1. Assume that $\succsim$ admits a WM representation $\langle u, D, \lambda \rangle$ with the properties in statement 2. To show that $\succsim$ satisfies A.1, A.3, A.4, A.5.1, and A.6, the same arguments for the necessity of the axioms in the proof of Theorem 1 can be applied since they rely on the same properties of u, D, and λ as we have here. Using the fact that for each $\varphi \in \mathbb{R}^S$, $\lambda(\varphi + t\mathbf{1})$ is non-increasing in t, it can be readily seen that $\succsim$ satisfies A.2.1.

Assume that $\succsim$ admits a VC representation $\langle u, K, \sigma \rangle$ with the properties in statement 3. To show that $\succsim$ satisfies A.1, A.3, A.4, and A.6, the same arguments for the necessity of the axioms in the proof of Theorem 2 can be applied since they rely on the same properties of u, K, and σ as we have here. Using the fact that for each $\varphi \in \mathbb{R}^S$, $\sigma(\varphi + t\mathbf{1})$ is non-increasing in t, it can be readily seen that $\succsim$ satisfies A.2.1. Using the fact that $\sigma \in \bar{\Sigma}(K)$, one can readily show that $\succsim$ satisfies A.5.1.

Assume that $\succsim$ admits an ambiguity averse representation $\langle u, G \rangle$ with the property in statement 4. By Theorems 3 and 5 of Cerreia-Vioglio et al. (2011b), $\succsim$ satisfies A.1, A.3, A.4, A.5.1, and A.6. Using the fact that for each $p \in \Delta$, $G(t, p) - t$ is non-decreasing in t, one can readily show that $\succsim$ satisfies A.2.1. □

Proof of Proposition 5 Let $\succsim$ satisfies A.1, A.2.1, A.3, A.4, A.5.1, and A.6. Let $\langle u, D, \lambda \rangle$ be a canonical WM representation of $\succsim$. It can be readily seen that A.2.1 implies A.2. Thus, by Theorem 1, D is the Bewley set, i.e., $\langle u, D \rangle$ is a Bewley representation of the unambiguous preference relation $\succsim^*$ induced from $\succsim$. Then by Proposition 11 of Cerreia-Vioglio et al. (2011b), $D = cl(co(\bigcup_{f \in \mathcal{F}} \pi(f)))$. To prove that for each $x \in X$, $D = cl(co(\bigcup_{f \sim x} \pi(f)))$, it suffices to prove that for each pair

$y, z \in X$, $cl(co(\bigcup_{f \sim y} \pi(f))) = cl(co(\bigcup_{f \sim z} \pi(f)))$. Fix $y, z \in X$. Assume without loss of generality that $y \succ z$.

Define $I : \mathbb{R}^S \to \mathbb{R}$ by setting for each $\varphi \in \mathbb{R}^S$, $I(\varphi) := \lambda(\varphi) \min_{p \in D} E_p \varphi + (1 - \lambda(\varphi)) \max_{p \in D} E_p \varphi$. By the similar arguments as in the proof of the necessity of the axioms in Proposition 1 and in the first paragraph of the proof of Theorem 1, I is non-decreasing, continuous, and quasi-concave. By the same arguments as in the second paragraph of the proof of Lemma 3, I is constant superadditive.

For each $x \in X$ and each $\varphi \in \mathbb{R}^S$, let $t_x(\varphi) \in \mathbb{R}$ be such that $I(\varphi - t_x(\varphi)\mathbf{1}) = u(x)$. Since I is constant superadditive, $t_x(\varphi)$ is uniquely determined. For each $x \in X$, define $\succsim_x$ on $\mathcal{F}$ by setting for each pair $f, g \in \mathcal{F}$,

$$f \succsim_x g \iff t_x(u(f)) \geq t_x(u(g)).$$

Let $\sim_x$ denote its indifference component.

Fix $x \in X$. The following four observations about $\sim_x$ will be useful. First, for each $x' \in X$, $t_x(u(x')) = u(x') - u(x)$, and in particular, $t_x(u(x)) = 0$. This follows immediately from the definitions of I and $t_x(u(x'))$. Second, for each $f \in \mathcal{F}$,

$$(i)\ x \succsim f \iff x \succsim_x f, \text{ and } (ii)\ x \sim f \iff x \sim_x f. \tag{49}$$

To see (i) of (49), note that

$$\begin{aligned} x \succsim f &\iff u(x) \geq I(u(f)) \iff 0 \geq t_x(u(f)) \iff t_x(u(x)) \\ &\geq t_x(u(f)) \iff x \succsim_x f, \end{aligned}$$

and by similar arguments, (ii) of (49) is true. Third, for each $\varphi \in \mathcal{F}$ and each $c \in \mathbb{R}$, $t_x(\varphi + c\mathbf{1}) = t_x(\varphi) + c$. This is because I is constant superadditive and $I(\varphi + c\mathbf{1} - t_x(\varphi + c\mathbf{1})\mathbf{1}) = u(x) = I(\varphi - t_x(\varphi)\mathbf{1}) = I(\varphi + c\mathbf{1} - (t_x(\varphi) + c)\mathbf{1})$. Fourth, $\succsim_x$ satisfies A.1–A.4, A.5.1, and A.6. By construction, $\succsim_x$ satisfies A.1. Since u is affine and onto, by the first observation, so is t_x when restricted to constant vectors, and thus, $\succsim_x$ satisfies A.2 and A.6. Since I is continuous and quasi-concave, by routine arguments, so is t_x, and thus $\succsim_x$ satisfies A.3 and A.5.1. Since I is constant superadditive, it can also be readily seen that $\succsim_x$ satisfies A.4.

For each $x \in X$, define $\pi_x : \mathcal{F} \rightrightarrows \Delta$ by setting for each $f \in \mathcal{F}$,

$$\pi_x(f) := \left\{ p \in \Delta : \sum_{s \in S} p_s f(s) \succsim_x \sum_{s \in S} p_s g(s) \text{ where } g \in \mathcal{F} \text{ implies } f \succsim_x g \right\}.$$

We claim that $\bigcup_{f \in \mathcal{F}} \pi_x(f) = \bigcup_{f \sim x} \pi(f)$. We first show that $\bigcup_{f \sim x} \pi(f) \subseteq \bigcup_{f \in \mathcal{F}} \pi_x(f)$. Let $x \in X$, $f \in \mathcal{F}$ with $f \sim x$, and $p \in \pi(f)$. We want to show that $p \in \pi_x(f)$. Let $g \in \mathcal{F}$ be such that $\sum_{s \in S} p_s f(s) \succsim_x \sum_{s \in S} p_s g(s)$. Then by the first observation in the previous paragraph, $\sum_{s \in S} p_s f(s) \succsim \sum_{s \in S} p_s g(s)$. Since $p \in \pi(f)$, $\sum_{s \in S} p_s f(s) \succsim \sum_{s \in S} p_s g(s)$ implies $f \succsim g$. Since $f \sim x$ and $f \succsim g$, $x \succsim g$. Since $f \sim x$ and $x \succsim g$, by applying the second observation in the previous paragraph twice, $f \succsim_x g$, as desired.

We now show that $\bigcup_{f\in\mathcal{F}} \pi_x(f) \subseteq \bigcup_{f\sim x} \pi(f)$. Let $x \in X$, $f \in \mathcal{F}$, and $p \in \pi_x(f)$. We need to show that there is $f' \in \mathcal{F}$ such that $f' \sim x$ and $p \in \pi(f')$. Since u is onto, we can pick $f' \in \mathcal{F}$ satisfying that

$$u(f') = u(f) - t_x(u(f))\mathbf{1}. \tag{50}$$

Since $I(u(f')) = I(u(f) - t_x(u(f))\mathbf{1}) = u(x)$, $f' \sim x$. To show that $p \in \pi(f')$, let $g' \in \mathcal{F}$ be such that $\sum_{s\in S} p_s f'(s) \succsim \sum_{s\in S} p_s g'(s)$, and we need to show that $f' \succsim g'$, or equivalently, $x \succsim g'$. Since u is onto, we can pick $g \in \mathcal{F}$ satisfying that

$$u(g) = u(g') + t_x(u(f))\mathbf{1}. \tag{51}$$

Since $\sum_{s\in S} p_s f'(s) \succsim \sum_{s\in S} p_s g'(s)$ and u is affine, $\sum_{s\in S} p_s u(f'(s)) \geq \sum_{s\in S} p_s u(g'(s))$. Thus, by (50) and (51), $\sum_{s\in S} p_s u(f(s)) \geq \sum_{s\in S} p_s u(g(s))$. Hence, $\sum_{s\in S} p_s f(s) \succsim \sum_{s\in S} p_s g(s)$. Then by the first observation, $\sum_{s\in S} p_s f(s) \succsim_x \sum_{s\in S} p_s g(s)$. Since $p \in \pi_x(f)$, $\sum_{s\in S} p_s f(s) \succsim_x \sum_{s\in S} p_s g(s)$ implies $f \succsim_x g$, i.e., $t_x(u(f)) \geq t_x(u(g))$. Since $t_x(u(f)) \geq t_x(u(g))$, by the third observation and by (50) and (51), $t_x(u(f')) \geq t_x(u(g'))$, i.e., $f' \succsim_x g'$. Since $f' \sim x$, by the second observation, $f' \sim_x x$. Thus, $x \succsim_x g'$, and again by the second observation, $x \succsim g'$, as desired.

Next, we claim that $\succsim_z$ is more ambiguity averse than $\succsim_y$. To show this, let $x \in X$ and $f \in \mathcal{F}$ be such that $x \succsim_y f$. We need to show that $x \succsim_z f$. Since $x \succsim_y f$, by the first observation, $u(x) - u(y) = t_y(u(x)) \geq t_y(u(f))$. Since $I(u(f) - t_y(u(f))\mathbf{1}) = u(y)$ and $u(x) - u(y) \geq t_y(u(f))$, and since I is non-decreasing, $I(u(f) - (u(x) - u(y))\mathbf{1}) \leq u(y)$. Thus, since I is constant superadditive and $y \succ z$,

$$\begin{aligned}
&I(u(f) - (u(x) - u(z))\mathbf{1}) \\
&\quad = I(u(f) - (u(x) - u(y))\mathbf{1} - (u(y) - u(z))\mathbf{1}) \\
&\quad \leq I(u(f) - (u(x) - u(y))\mathbf{1}) - (u(y) - u(z)) \\
&\quad \leq u(y) - (u(y) - u(z)) = u(z).
\end{aligned}$$

Since $I(u(f) - (u(x) - u(z))\mathbf{1}) \leq u(z) = I(u(f) - t_z(u(f))\mathbf{1})$, and since I is constant superadditive, $u(x) - u(z) \geq t_z(u(f))$. Thus, by the first observation, $t_z(u(x)) = u(x) - u(z) \geq t_z(u(f))$, i.e., $x \succsim_z f$, as desired.

Since $\succsim_z$ is more ambiguity averse than $\succsim_y$, by Propositions 6 and 11 and Theorem 10 of Cerreia-Vioglio et al. (2011b), and by the first observation, $cl(co(\bigcup_{f\in\mathcal{F}} \pi_y(f))) \subseteq cl(co(\bigcup_{f\in\mathcal{F}} \pi_z(f)))$.[30] We have shown that for each $x \in X$, $\bigcup_{f\in\mathcal{F}} \pi_x(f) = \bigcup_{f\sim x} \pi(f)$. Thus, $cl(co(\bigcup_{f\sim y} \pi(f))) \subseteq cl(co(\bigcup_{f\sim z} \pi(f)))$.

---

30 Under A.1–A.4, the definition of a more ambiguity averse preference adopted in this paper is equivalent to that adopted by Cerreia-Vioglio et al. (2011b).

Conversely, we show that $cl(co(\bigcup_{f\sim z}\pi(f))) \subseteq cl(co(\bigcup_{f\sim y}\pi(f)))$. It is sufficient to show that $\bigcup_{f\sim z}\pi(f) \subseteq cl(co(\bigcup_{f\sim y}\pi(f)))$. Suppose to the contrary that there exist $g \in \mathcal{F}$ with $g \sim z$ and $q \in \pi(g)\backslash cl(co(\bigcup_{f\sim y}\pi(f)))$. Then by a standard separation theorem, there is $\varphi \in \mathbb{R}^S$ such that for each $p \in cl(co(\bigcup_{f\sim y}\pi(f)))$, $E_q\varphi < 0 < E_p\varphi$. Pick $n \in N$ such that $E_q n\varphi < E_q u(g)$ and for each $p \in cl(co(\bigcup_{f\sim y}\pi(f)))$, $E_p n\varphi > u(y)$. Let $h \in \mathcal{F}$ be such that $u(h) = n\varphi$. Since $E_q u(g) > E_q n\varphi = E_q u(h)$, and since $q \in \pi(g)$, $g \succsim h$. Since $g \sim z$ and $g \succsim h$, $z \succsim h$. By the fourth observation, $\succsim_y$ satisfies A.1–A.4, A.5.1, and A.6. Thus, by the first observation and by applying Proposition 11 of Cerreia-Vioglio et al. (2011b) for $\succsim_y$, if for each $p \in cl(co(\bigcup_{f\in\mathcal{F}}\pi_y(f)))$, $E_p u(h) \geq u(y)$, then $h \succsim^*_y y$, where $\succsim^*$ is the unambiguous preference relation induced from $\succsim_y$. Since $cl(co(\bigcup_{f\in\mathcal{F}}\pi_y(f))) = cl(co(\bigcup_{f\sim y}\pi(f)))$, and since for each $p \in cl(co(\bigcup_{f\sim y}\pi(f)))$, $E_p u(h) = E_p n\varphi > u(y)$, $h \succsim^*_y y$, which in particular implies that $h \succsim y$. Since $h \succsim y$ and $y \succ z$, $h \succ z$, which contradicts that $z \succsim h$, as desired. □

Proof of Proposition 6 Let $\succsim$ satisfy A.1–A.4, and A.6. We first prove the sufficiency of A.2.3. Assume that $\succsim$ also satisfies A.2.3. By Lemma 1, there exist an affine onto utility function $u : X \to \mathbb{R}$ and a normalized, monotone, and continuous functional $I : \mathbb{R}^S \to \mathbb{R}$ satisfying (22).

We claim that I is constant additive. Since $\succsim$ satisfies A.2.3, by the similar arguments as in the second paragraph of the proof of Lemma 3, for each $\varphi \in \mathbb{R}^S$ and each $t \in \mathbb{R}$, $I(\varphi + t\mathbf{1}) \geq I(\varphi) + t$, and hence $I((\varphi + t\mathbf{1}) - t\mathbf{1}) \geq I(\varphi + t\mathbf{1}) - t$. Then for each $\varphi \in \mathbb{R}^S$ and each $t \in \mathbb{R}$,

$$I(\varphi + t\mathbf{1}) \geq I(\varphi) + t = I(\varphi + t\mathbf{1} - t\mathbf{1}) + t \geq I(\varphi + t\mathbf{1}),$$

which implies that $I(\varphi + t\mathbf{1}) = I(\varphi) + t$, as desired.

Let $\alpha \in (0, 1)$, $x, y \in X$, and $f, g \in \mathcal{F}$ be such that $f \trianglerighteq g$ and $\alpha f + (1-\alpha)x \succsim \alpha g + (1-\alpha)x$. Thus, since u is affine and I satisfies 22, $I(\alpha u(f) + (1-\alpha)u(x)) \geq I(\alpha u(g) + (1-\alpha)u(x))$. Then since I is constant additive,

$$\begin{aligned}
&I(\alpha u(f) + (1-\alpha)u(y)) \\
&\quad = I(\alpha u(f) + (1-\alpha)u(x)) + (1-\alpha)(u(y) - u(x)) \\
&\quad \geq I(\alpha u(g) + (1-\alpha)u(x)) + (1-\alpha)(u(y) - u(x)) \\
&\quad = I(\alpha u(g) + (1-\alpha)u(y)).
\end{aligned}$$

Hence, $\alpha f + (1-\alpha)y \succsim \alpha g + (1-\alpha)y$, as desired.

The necessity of A.2.3 follows from the fact that for each $f \in \mathcal{F}$ and $z \in X$, $f \trianglerighteq z$, which can be readily seen by definition. □

Proof of Proposition 7 Let $u : \mathbb{R} \to \mathbb{R}$ be an affine onto function and $\phi : \mathbb{R} \to \mathbb{R}$ an increasing, concave, and twice differentiable function. For each countably additive Borel probability measure μ on Δ, define $I_\mu : \mathbb{R}^S \to \mathbb{R}$ by setting for each $\varphi \in \mathbb{R}$,

$$I_\mu(\varphi) := \phi^{-1}\left(\int_{p\in\Delta} \phi(E_p\varphi)d\mu(p)\right).$$

By definition, I is normalized.

To show the "if" direction, assume that $-\frac{\phi''}{\phi'}$ is non-increasing. Let $\succsim$ admit a smooth ambiguity representation $\langle u, \phi, \mu\rangle$ for some countably additive Borel probability measure μ on Δ. We want to show that $\succsim$ satisfies A.2.1. Since $\langle u, \phi, \mu\rangle$ represents $\succsim$, I satisfies (22). By the analogous version of Lemma 52 in Cerreia-Vioglio et al. (2011b), and by the analogous arguments as in the proof of their Proposition 53, I is constant superadditive. It can be readily seen that since I satisfies (22), and since I is normalized and constant superadditive, $\succsim$ satisfies A.2.1.

Conversely, assume that each preference relation that admits a smooth ambiguity representation $\langle u, \phi, \mu\rangle$ for some countably additive Borel probability measure μ on Δ satisfies A.2.1. We want to show that $-\frac{\phi''}{\phi'}$ is non-increasing. By the analogous version of Lemma 52 in Cerreia-Vioglio et al. (2011b), it is equivalent to show that for each $t \in \mathbb{R}_+$, $J_t : \phi(\mathbb{R}) \to \mathbb{R}$, defined by setting for each $c \in \phi(\mathbb{R})$,

$$J_t(c) = \phi[\phi^{-1}(c) + t],$$

is convex on $\phi(\mathbb{R})$. Note that since ϕ is increasing, ϕ^{-1} is well-defined. Let $c, c' \in \phi(\mathbb{R})$ and $\alpha \in [0, 1]$. Let $x, x' \in X$ be such that $\phi(u(x)) = c$ and $\phi(u(x')) = c'$. Let $s, s' \in S$ and $f \in \mathcal{F}$ be such that $f(s) = x$ and $f(s') = x'$. Let δ_s and $\delta_{s'}$ be the degenerate probability measures which assign probability one, respectively, to s and s'. Let μ be a probability measure on Δ which assigns probability α to δ_s and $1 - \alpha$ to $\delta_{s'}$. Let $\succsim$ admit the smooth ambiguity representation $\langle u, \phi, \mu\rangle$. Then $\succsim$ satisfies A.2.1. By the similar arguments as in the proof of Lemma 3, I_μ is constant superadditive. Then for each $t \in \mathbb{R}_+$ and each $\alpha \in [0, 1]$,

$$\begin{aligned}
&J_t(\alpha c + (1-\alpha)c') = J_t(\alpha\phi(u(f(s))) + (1-\alpha)\phi(u(f(s'))))\\
&= J_t\left(\int_{p\in\Delta} \phi(E_p u(f))d\mu(p)\right) = \phi[I_\mu(u(f)) + t] \le \phi[I_\mu(u(f) + t\mathbf{1})]\\
&= \int \phi(E_p(u(f) + t\mathbf{1}))d\mu(p) = \alpha\phi(u(x) + t) + (1-\alpha)\phi(u(x') + t)\\
&= \alpha\phi[\phi^{-1}(c) + t] + (1-\alpha)\phi[\phi^{-1}(c') + t] = \alpha J_t(c) + (1-\alpha)J_t(c'),
\end{aligned}$$

as desired.

Similarly, it can be shown that each preference relation that admits a smooth ambiguity representation $\langle u, \phi, \mu \rangle$ for some countably additive Borel probability measure μ on Δ satisfies A.2.2 if and only if $-\frac{\phi''}{\phi'}$ is non-decreasing. $\square$

References

Anscombe, F., Aumann, R.: A definition of subjective probability. Ann. Math. Stat. **34**, 199–205 (1963)

Arrow, K.: Aspects of the Theory of Risk-Bearing. Academic Bookstore, Helsinki (1965)

Bewley, T.F.: Knightian decision theory: Part I. Decis. Econ. Finan. **25**(2), 79–110 (2002)

Baillon, A., Placido, L.: Testing constant absolute and relative ambiguity aversion, mimeo (2015)

Cerreia-Vioglio, S., Ghirardato, P., Maccheroni, F., Marinacci, M., Siniscalchi, M.: Rational preferences under ambiguity. Econ. Theory **48**(2), 341–375 (2011). https://doi.org/10.1007/s00199-011-0643-5

Cerreia-Vioglio, S., Maccheroni, F., Marinacci, M., Montrucchio, L.: Ambiguity averse preferences. J. Econ. Theory **146**(4), 1275–1330 (2011)

Cerreia-Vioglio, S., Maccheroni, F., Marinacci M., Montrucchio, L.: Absolute and relative ambiguity aversion: a preferential approach, mimeo (2017)

Chambers, R., Grant, S., Polar, B., Quiggin, J.: A two-parameter model of dispersion aversion. J. Econ. Theory **150**, 611–641 (2014)

Chateauneuf, A., Faro, J.H.: Ambiguity through confidence functions. J. Math. Econ. **45**, 535–558 (2009)

Cherbonniera, F., Gollier, C.: Decreasing aversion under ambiguity. J. Econ. Theory **157**, 606–623 (2015)

Dupuis, P., Ellis, R.S.: A weak convergence approach to the theory of large deviations. Wiley, New York (1997)

Ellsberg, D.: Risk, ambiguity and the Savage Axioms. Q. J. Econ. **75**, 643–669 (1961)

Gajdos, T., Hayashi, T., Tallon, J.-M., Vergnaud, J.-C.: Attitude toward imprecise information. J. Econ. Theory **140**(1), 27–65 (2008)

Ghirardato, P., Marinacci, M.: Ambiguity made precise: a comparative foundation. J. Econ. Theory **102**, 251–289 (2002)

Ghirardato, P., Maccheroni, F., Marinacci, M.: Differentiating ambiguity and ambiguity attitude. J. Econ. Theory **118**, 133–173 (2004)

Gilboa, I., Schmeidler, D.: Maximin expected utility with non-unique prior. J. Math. Econ. **18**, 141–153 (1989)

Grant, S., Polak, B.: Mean-dispersion preferences and constant absolute ambiguity aversion. J. Econ. Theory **148**(4), 1361–1398 (2013)

Hansen, L., Sargent, T.: Robust control and model ambiguity. Am. Econ. Rev. **91**(2), 60–66 (2001)

Hansen, L., Sargent, T.: Robustness. Princeton University Press, Princeton (2008)

Hill, B.: Confidence and decision. Games Econ. Behav. **82**, 675–692 (2013)

Hurwicz, L.: Some specification problems and applications to econometric methods (Abstract). Econometrica **19**, 343–344 (1951)

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

Knight, F.: Risk, Ambiguity and Profit. Houghton Mifflin, Boston (1921)

Kopylov, I.: Procedural Rationality in the Multiple Prior Model. Mimeo. University of Rochester, Rochester (2001)

Kopylov, I.: Choice deferral and ambiguity aversion. Theor. Econ. **4**, 199–225 (2009)

Lang, M.: First-order and second-order ambiguity aversion. Manag. Sci. **63**(4), 901–1269 (2017)

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

Nehring, K.: Bernoulli without Bayes: a theory of utility-sophisticated preferences under ambiguity, mimeo (2007)

Olszewski, W.: Preferences over sets of lotteries. Rev. Econ. Stud. **74**, 567–595 (2007)

Pratt, J.: Risk aversion in the small and in the large. Econometrica **32**(1/2), 122–136 (1964)

Rigotti, Shannon, Strazalecki, J.: Subjective beliefs and ex-ante trade. Econometrica **76**(5), 1167–1190 (2008)

Schmeidler, D.: Subjective probability and expected utility without additivity. Econometrica **57**, 571–587 (1989)
Strzalecki, T.: Axiomatic foundations of multiplier preferences. Econometrica **79**(1), 47–73 (2011a)
Strzalecki, T.: Probabilistic sophistication and variational preferences. J. Econ. Theory **146**, 2117–2125 (2011b)
Wald, A.: Basic ideas of a general theory of statistical decisions rules. In: Wald, A. (ed.) Selected Papers in Statistics and Probability, pp. 656–668 (1950a)
Wald, A.: Statistical Decision Functions. Wiley, New York (1950b)
Xue, J.: Three representations of preferences with decreasing absolute uncertainty aversion, mimeo (2012)