Comparative Ambiguity Attitudes*

Fan Wang†

Abstract

I propose a new behavioral definition of comparative ambiguity aversion, which, in contrast to standard definitions, establishes a separation between ambiguity attitude and risk attitude. Ambiguity attitudes are compared using a notion called *matching probability*, which reflects how an agent reduces ambiguity to risk. My definition enables well-accepted formal characterizations without requiring the agents under comparison to have the same von Neumann-Morgenstern utility; results are established for a wide range of models, including, in particular, maxmin expected utility (Gilboa and Schmeidler 1989) and the smooth ambiguity model (Klibanoff, Marinacci, and Mukerji 2005). These results naturally point to a decomposition of differences in willingness to pay into a factor attributable to differences in ambiguity attitude, and another to risk attitude. The decomposition is easily implementable and it is free of structural or parametric assumptions. I also use the main idea to formulate and characterize absolute and relative ambiguity attitude in isolation from risk attitude. This paper therefore expands the scope of comparative ambiguity attitudes, which is fundamentally useful for theoretical results as well as for empirical tests.

*Date: November 14, 2019.

†HEC, Paris. Email: fan.wang1@hec.edu. I am grateful to Mohammed Abdellaoui, Itzhak Gilboa, and Stefania Minardi for their continuous encouragement and many useful discussions. I thank Massimo Marinacci, Sujoy Mukerji, Peter Wakker, and other participants of D-TEA workshop 2019 for their feedbacks.

1 Introduction

Ever since ambiguity aversion was formally modeled, researchers have been interested in defining and characterizing comparative ambiguity attitudes. What does it mean for one agent to be "more ambiguity averse" than another? How is such a relation reflected in the utility representations of the two individuals' preferences? And, most importantly, what are the economic implications of such a comparison? These questions parallel the questions asked about the "more risk averse" relation. Clearly, such a relation is not expected to be complete; one does not assume that any two decision makers (DMs) can be compared so that one is more risk or ambiguity averse than the other. Yet, when the relation does hold, one could hope for a variety of insightful results, as in the case of risk attitudes (see Pratt 1964, Arrow 1965, Yaari 1969, Ross 1981, Kihlstrom, Roth, and Schmeidler 1981).

The literature on comparative ambiguity attitudes has indeed proposed several insightful behavioral conditions that define what it means for one decision maker to be more ambiguity averse than another. However, these definitions often assume that the DMs under comparison have the same preference over lotteries (or consequences, or ambiguity-free acts in general).[1] In the popular Anscombe and Aumann (1963) framework, this amounts to assuming the same (cardinal) von Neumann-Morgenstern (vNM) utility function, and therefore, the same risk attitude. Restricting attention to pairs of DMs who share preferences over (risky, but not ambiguous) lotteries is an insightful simplification. Importantly, it guarantees that, whatever implications the comparison might have in terms of economic behavior would be a result of differential ambiguity attitudes rather than, say, of risk attitudes. Thus, most of the existing definitions allow a clear identification of the effect of ambiguity aversion per se, purified of any other personal differences DMs might have, whether ordinal (under certainty) or cardinal (under risk).

However, this simplification comes at a cost. First, it restricts the scope of theoretical results one can hope to derive from comparisons of ambiguity attitudes. For example, suppose that a venture capitalist (VC) seeks to invest in a start-up project owned by an entrepreneur, whose chance of success the VC perceives as being ambiguous (Knight 1921). The former, investing in many firms, may be more tolerant toward risk than is the latter. By contrast, the entrepreneur tends to know the project better, and therefore, in decision-theoretic language, may be more tolerant toward ambiguity (when applied to the start-up project) than is the VC. Thus one can consider the optimal contract between the two parties, one of whom is

[1]See, for example, Epstein (1999), Epstein and Zhang (2001), Ghirardato and Marinacci (2002), Klibanoff, Marinacci and Mukerji (2005), Maccheroni, Marinacci, and Rustichini (2006), Chateauneuf and Faro (2009), Cerreia-Vioglio, Maccheroni, Marinacci, and Montrucchio (2011b), and Cerreia-Vioglio, Maccheroni, and Marinacci (2019).

more risk averse than the other, while the latter is more ambiguity averse than the former (see Example 7 for a setup). Whatever results one might hope to establish could not be stated using the most popular definitions of comparative ambiguity attitudes. To consider another example, suppose that we believe that DM 1 would invest less in the Dow than DM 2, as compared to their respective investments in Treasury Bills, if DM 1 is more risk averse or more ambiguity averse than DM 2, *or both* (see the two-asset example in section 5.2). Existing definitions of "more ambiguity averse" render this conjecture meaningless: they allow for the conjecture that *either* risk aversion *or* ambiguity aversion might yield the desired conclusion, but not both because, by definition, ambiguity attitudes cannot be compared across individuals whose risk attitudes differ.

Second, the restricted domain of these definitions can also be a hurdle to empirical work. Assume, for example, that a researcher wishes to establish that old people tend to be more ambiguity averse than young people. To this end, one might sample DMs of various ages, confront them with incentivized financial decisions and compare their decisions. However, the current definitions would not endorse any conclusion about ambiguity attitudes unless one could show that the DMs in question had the same ordinal preferences and the same risk attitudes. As this might be hard to establish, comparisons of ambiguity attitudes would be very difficult to conduct.[2]

Third, the restriction prevents one from contrasting the economic impact of ambiguity attitude against that of risk attitude, a problem that is especially acute when their impacts are opposite. Take, for example, when an agent needs to pay efforts in order to reduce the probability of bad outcomes. Risk aversion is likely to induce the agent to make a high effort. Ambiguity aversion, on the other hand, makes the agent not value moderate effort too highly, because, having reduced some risks, the agent tends to focus on the remaining ones. As effort is costly, the agent is likely to be driven by ambiguity aversion to take extreme actions, that is, either low effort or high effort but not moderate effort (see section 5.1 for a rigorous analysis).

This paper suggests a new formulation of the relation "more ambiguity averse", which applies also to pairs of preference relations that do not agree on (objective) lotteries, or even on sure outcomes. The key insight is that, instead of following classical formulations to compare the two DMs' willingness to pay for a common (ambiguous) act, which necessarily gets both risk attitude and ambiguity attitude involved, we compare how the two DMs reduce ambiguity to risk. We argue that our formulation captures an intuitive notion of "pure" ambiguity attitudes, and that the comparison it is based on is not contaminated by differential risk attitudes or by varying tastes more generally. This is especially desirable

[2]Indeed, Deakin et al. (2004) and Dohmen et al. (2011) show that older people tend to be less risk averse.

since many models of decision under uncertainty already use different formal components to model ambiguity attitude and risk attitude separately. As a consequence, we are naturally led to a behavioral decomposition of two DMs' difference in willingness to pay for any act into a part attributable to their difference in ambiguity attitude and a part attributable to their difference in risk attitude. The idea also suggests a new formulation and a new characterization of absolute and relative ambiguity attitudes, which provide decision-theoretic guidance to the modeling of ambiguity attitude and risk attitude that can change with wealth simultaneously.

For a rough understanding of the main idea, let us revisit the Ellsberg's (1961) two-urn example. The "known" urn contains 50 black balls and 50 red balls, while the "unknown" urn contains 100 balls with an unknown composition of black and red balls. The experimenter announces a color that is either black or red and then asks the DM to choose a urn to bet on. A ball is then randomly drawn from the chosen urn and the DM receives \$100 if the ball has the announced color and \$0 if not. The DM is *ambiguity averse* if he or she always chooses to bet on the known urn no matter what color is announced. To compare two DMs' ambiguity attitudes, one method is to compare their willingness to pay for the two possible bets on the unknown urn, and the DM with lower willingness to pay is deemed more ambiguity averse. This method is implicitly used in the classical formulations of comparative ambiguity attitudes of Epstein (1999) and Ghirardato and Marinacci (2002), and, due to the theoretical justifications given by the two papers, it has been widely adopted by later theoretical works. However, a major limitation of this method is that it requires the two DMs to share the same utility function, because a DM might have lower willingness to pay simply due to a more concave utility function. Another method is to first elicit, for each DM $i \in \{1, 2\}$ and each color $k \in \{black, red\}$, the percentage number $n_{i,k}$ such that when color k is announced the DM is indifferent between betting on the unknown urn and betting on another known urn with $n_{i,k}$ percentage of balls in color k, and then, the DM with lower elicited percentage numbers is deemed more ambiguity averse. This method does not need the common utility assumption and it is essentially adopted by Baillon, Huang, Selim, and Wakker (2018) and Dimmock, Kouwenberg, and Wakker (2016) for empirical work. However, a general theoretical justification for this method is missing. The first part of this paper gives a new formulation of comparative ambiguity attitudes based on this method and provides a general decision-theoretic justification (see section 2 and the complementary section 6), which show that the method not only works under a large range of formal models but also follows the same fundamental idea pioneered by Ghirardato and Marinacci (2002) in their classical formulation.

More formally, our starting point is the observation that in the settings of many axiomatic models of decisions under ambiguity, and certainly within the reach of experimental design, we can find some component other than money that is commonly desirable to all decision makers. Take for example the widely used Anscombe and Aumann (1963) setting, which involves both risk and ambiguity. In this setting, whenever a DM prefers prize a to prize b, then for any lottery over $\{a, b\}$, the DM likes it better if the lottery puts a higher probability on prize a. Notice that this desire for higher probabilities on the subjectively better prize is common to all DMs, even if they have ordinally different preferences over prizes. In a more general setting, one can have a state space as in Savage (1954) and a convex set of consequences (see, for example, Maccheroni, Marinacci, and Rustichini 2006), a higher coefficient $\alpha \in [0, 1]$ associated with the subjectively better consequence in a mixture between two consequences is again commonly desirable to all DMs.

A commonly desirable component as such might naturally serve as a reference system for comparing ambiguity attitudes, leaving full flexibility to preferences. To illustrates this point, imagine that DM $i \in \{1, 2\}$ has a complete and transitive preference $\succsim^i$ over acts. For any probability $p \in [0, 1]$, $(a, p; b, 1-p)$ denotes the lottery that yields prize a with probability p and prize b with probability $1-p$. For any event E, $(a, E; b, E^c)$ denotes the act that yields prize a in event E and prize b in event E^c. Fix four prizes $a \succ^1 b$ and $c \succ^2 d$ and for any event E ask DM $i \in \{1, 2\}$ to find the *matching probability* $p^i_E \in [0, 1]$ such that

$$\begin{aligned} (a, E; b, E^c) &\sim^1 (a, p^1_E; b, 1-p^1_E), \\ (c, E; d, E^c) &\sim^2 (c, p^2_E; d, 1-p^2_E). \end{aligned}$$

Our intuition is that DM1 is at least as ambiguity averse as DM2 if

$$p^1_E \leq p^2_E, \ \textit{for all event } E. \tag{1.1}$$

The reason is as follows. Suppose $\mathcal{S}$ denotes the whole state space. As the sure-win act $(a, \mathcal{S}; b, \emptyset)$ should get matched by $p^1_{\mathcal{S}} = 100\%$ chance of winning prize a, the difference $1 - p^1_E$ reflects DM 1's disutility derived from a smaller event of winning prize a and the ambiguity associated with it. If such disutility of DM 1 is uniformly higher than that of DM 2 across *all* events, it is likely that person 1 derives more disutility from ambiguity.

Importantly, in testing this condition, each decision maker is asked to compare only acts with two possible prizes at a time (a, b for DM 1 and c, d for DM 2). Thus, risk attitudes play no role in the comparisons used in the definition. This fact has two related implications. First, the comparison we obtain is "pure" in the sense that it cannot be contaminated by differential risk attitudes. Second, it allows us to compare individuals whose risk attitudes differ. Moreover, as the prizes can vary from one individual to another, the individuals need not agree on the ordinal ranking of the prizes in order for the definition to apply.

Let us compare this condition with existing ones. Under mild technical assumptions,[3] statement (1.1) is equivalent to

$$(a, E; b, E^c) \succcurlyeq^1 (a, p; b, 1-p) \implies (c, E; d, E^c) \succcurlyeq^2 (c, p; d, 1-p), \tag{1.2}$$

for any event E and any probability $p \in [0, 1]$. When the two DMs do share the same ordinal preference over prizes, we can set $a = c$ and $b = d$, and the above condition is easily seen to be implied by Ghirardato and Marinacci's (2002) popular formulation of comparative ambiguity attitudes, according to which DM 1 is more ambiguity averse than DM 2 if and only if for any act f and any lottery x,

$$f \succcurlyeq^1 x \implies f \succcurlyeq^2 x. \tag{1.3}$$

However, condition (1.3) implies that the DMs share the same preference over lotteries, which is then used as a common reference system to indicate that DM 1 always has a weakly lower willingness to pay for ambiguous acts than DM 2. In comparison, the innovation in condition (1.2) is twofold. First, instead of using willingness to pay, we use matching probabilities (or, in general, probabilities on the subjectively better prize) as the reference system. Second, we restrict their condition to acts and lotteries that involve only two fixed prizes, and also generalize it by allowing the subjectively better prize to vary across the DMs.

As a simple application, we note that when the DMs have maxmin preferences (Gilboa and Schmeidler 1989) and there are only two states of the world, condition (1.2) is equivalent to DM 1 having a larger set of priors, and this equivalence will be proven more generally in Theorem 3 below. Many existing definitions of "more ambiguity averse" do indeed correspond to set inclusion between the sets of priors. Indeed, if the model captures ambiguity aversion by using a set of priors rather than a single one, it seems natural that a larger set, in terms of set inclusion, would capture "more ambiguity averse". But, as mentioned above, existing definitions can only apply this comparison to DMs who share a utility function, whereas our definition does not.

The main message of this paper is that, under the expected utility hypothesis over risk, comparing matching probabilities of a suitable set of acts amounts to comparing ambiguity attitudes independent of risk attitudes, while obtaining identical characterizations in terms of the formal representation of ambiguity attitude. In other words, if one defines the "more ambiguity averse" relation in terms of the utility representation and asks what does this definition mean behaviorally, the present contribution provides a characterization that is applicable in a significantly wider set of conditions. The theoretical advantage of this extension

[3] For example, $\succcurlyeq^i$ is continuous, monotone, and respects first order stochastic dominance over risk.

of the domain is that one can study the implications of comparative ambiguity attitudes in much richer setups than hitherto allowed, and to contrast risk and ambiguity aversion. From a more practical point of view, this definition allows to test the "more ambiguity averse" condition without measuring risk aversion at all, and without restricting attention to the cases where ordinal preferences and risk aversion are the same for the individuals compared. We establish formal characterizations of the "more ambiguity averse" relation for maxmin preferences (Theorem 3), Bewley preferences (Corollary 4), the smooth ambiguity model (Theorem 6), Choquet expected utility (Proposition 11), the generalized α-maxmin model (Proposition 12), homothetic preferences (Proposition 13), the general class of uncertainty averse preferences (Theorem 14), variational preferences (Corollary 15), and multiplier preferences (Corollary 16).

As these results justify the view that ambiguity attitude is captured by matching probabilities, we are naturally led to a behavioral decomposition of two DMs' difference willingness to pay for any act into two parts, one attributable to their difference in ambiguity attitude, and the other -- in risk attitude. The key step for this decomposition is to ask the two DMs to switch their matching probabilities, which amounts to having the DMs switch their ambiguity attitude but retain their own risk attitude, and then re-evaluate their willingness to pay accordingly. This decomposition is easily implementable and it is free of structural or parametric assumptions, thereby providing a simple set-up within which the venture capitalist-entrepreneur problem can be analyzed (see Example 7).

The main idea also sheds light on the search of a behavioral formulation and a formal characterization of absolute and relative ambiguity attitude, by treating the same DM at different wealth levels as distinct DMs. Existing results that use Ghirardato and Marinacci's (2002) comparative notion (see Cerreia-Vioglio, Maccheroni, and Marinacci 2019) require the DM to have constant absolute (or relative) risk attitude. This restriction is circumvented by our results, making it possible to study the wealth effects through the risk channel and the ambiguity channel simultaneously.

The paper is organized as follows. Section 2 proposes our formulation of comparative ambiguity attitudes and applies it to maxmin expected utility, Bewley preferences, and the smooth ambiguity model. Section 3 provides a behavioral decomposition of two DMs' difference in willingness to pay for any act into a risk component and an ambiguity component. Section 4 applies our main idea to the formulation and characterization of absolute and relative ambiguity attitudes. Section 5 discusses how the tools developed here might help economists better understand economic phenomena at large. Section 6 complements Section 2 by applying the formulation to a collection of other popular models of decision under uncertainty. Section 7 concludes by discussing technical issues surrounding the formulation.

1.1 Related Literature

A sizable body of works focus on the formulation and the characterization of comparative ambiguity attitudes, starting from Schmeidler's (1989) seminal paper. The literature mostly settles on the formulation of Ghirardato and Marinacci (2002) under expected utility over risk (and the closely related condition of Epstein (1999) under general risk preferences), which mimics classical comparative risk attitudes. This paper joins this line of inquiry; our behavioral formulation generalizes Ghirardato and Marinacci's (2002) in a small domain to free it from restricting other aspects of behaviors.

The key insight of our formulation is rooted in the likelihood approach to decisions under uncertainty. This approach was first proposed by Gilboa (1987) and was further developed by Abdellaoui and Wakker (2005) in the Savagean framework, which involves ambiguity only. Abdellaoui (2002) developed this approach in the risk domain. This paper and Wang (2019) extend this approach to models where both ambiguity and risk are present, by using well-defined uncertainty (i.e., risk) as a device to measure less well-defined uncertainty (i.e., ambiguity) and other aspects of decision making.

This paper is most closely related to works that use matching probabilities to measure ambiguity aversion (Baillon, Bleichrodt, Li, and Wakker 2019, Baillon, Huang, Selim, and Wakker 2018, and Dimmock, Kouwenberg, and Wakker 2016). These works aim at providing guidelines to empirical executions of comparisons of ambiguity attitudes. To this end, they propose indices that aggregate, in various ways, matching probabilities of simple binary acts, and focus on ingenious experimental designs that allow these indices to reflect ambiguity attitudes. In contrast, this paper takes one step back and provides the general decision theoretic foundation for the likelihood approach to comparative ambiguity attitudes.

Some previous works compared concepts related to ambiguity attitude without a common-utility restriction. See, for example, Gajdos, Hayashi, Tallon, and Vergnaud (2008), which compare a notion of aversion to imprecise information by comparing willingness to bet on imprecise information, and Kopylov (2016), which compare a notion of confidence by comparing behaviors that have a taste of hedging. A technically similar condition was independently used by Hill (2019) to isolate state-dependent utility from perceived ambiguity.

Finally, on absolute and relative ambiguity attitudes, this paper is most closely related to Cerreia-Vioglio, Maccheroni, and Marinacci (2019). In particular, our behavioral conditions do not impose constant absolute (or relative) risk attitude on the DM. In support of Baillon and Placido's (2019) empirical evidence, we give concrete examples of decreasing absolute ambiguity attitudes under a few formal models.

2 A Formulation of Comparative Ambiguity Attitudes

For an intuitive behavioral interpretation, we adopt the two-stage Anscombe and Aumann (1963) framework throughout the paper.[4] The primitives are as follows. There is a set $\mathcal{S}$ of *states* of the world and a σ-algebra Σ of subsets of $\mathcal{S}$ that we call *events*. There is a set $\mathcal{X}$ of *consequences*, each of which is a *lottery* (i.e., a probability distribution with finite support) over a set $\mathcal{M}$ of *prizes*. An *act* is a function $f : \mathcal{S} \to \mathcal{X}$ measurable with respect to Σ that takes finitely many values. We denote by $\mathcal{F}$ the collection of all acts.

Following convention, any prize $m \in \mathcal{M}$ is also understood as a degenerate lottery, and any lottery $x \in \mathcal{X}$ is also understood as a constant act, where all states are mapped to the same lottery. For any lotteries $x, y \in X$ and any $\alpha \in [0,1]$, the *lottery mixture* $\alpha x + (1-\alpha)y$ stands for the lottery that assigns probability $\alpha x(m) + (1-\alpha)y(m)$ to any prize $m \in \mathcal{M}$. For any two prizes $a, b \in \mathcal{M}$ and any $\alpha \in [0,1]$, we use $(a, \alpha; b, 1-\alpha)$ to denote the lottery that yields a with probability α and b with probability $1-\alpha$. Specific to this paper, we introduce the notation $\mathcal{F}_{a,b}$ (for any prizes $a, b \in \mathcal{M}$) to denote the family of acts that map states to the "lottery line segment" connecting a to b, that is, $\mathcal{F}_{a,b} = \{f \in \mathcal{F} : \forall s \in \mathcal{S}, \exists \alpha \in [0,1], f(s) = (a, \alpha; b, 1-\alpha)\}$.

Decision maker (DM) $i \in \{1, 2\}$ has a preference over $\mathcal{F}$, which is modeled as a complete and transitive order $\succcurlyeq^i$ on $\mathcal{F}$. The asymmetric and symmetric part of $\succcurlyeq^i$ are denoted by $\succ^i$ and $\sim^i$ respectively, with $\succ^i$ assumed nonempty.

As $\succcurlyeq^1$ and $\succcurlyeq^2$ may not even agree on sure prizes, we first describe in the following definition when two acts are comparable between the DMs.

Definition 1. For *any* four prizes $a \succ^1 b$ and $c \succ^2 d$, acts $f \in \mathcal{F}_{a,b}$ and $g \in \mathcal{F}_{c,d}$ are $(\succcurlyeq^1, \succcurlyeq^2)$-*aligned*, if $f(s)(a) = g(s)(c)$ for all $s \in \mathcal{S}$.

When two acts $f \in \mathcal{F}_{a,b}$ and $g \in \mathcal{F}_{c,d}$ are $(\succcurlyeq^1, \succcurlyeq^2)$-aligned, on any state s lotteries $f(s)$ and $g(s)$ share the same probability associated with the subjectively better prize. This suggests that f represents a "profile of winning-probabilities" among $\mathcal{F}_{a,b}$ as strong to DM 1 as g's profile among $\mathcal{F}_{c,d}$ is to DM 2. Inherent in this definition is a change of perspective from prizes to probabilities: instead of finding prizes or acts that are commonly desirable, we exploit the probabilities on subjective better prizes. To streamline exposition, we will simply use the terminology "aligned acts", with the preference pair $(\succcurlyeq^1, \succcurlyeq^2)$ invoked implicitly.[5]

[4] All our results also hold in a more general setting where the set of consequences is a convex subset of a real vector space (see, for example, Klibanoff, Marinacci, and Mukerji 2005, and Maccheroni, Marinacci, and Rustichini 2006), by re-interpreting our notion of lottery mixture through the operations of multiplication and addition inherent to the vector space. Settings with even weaker structures are discussed in section 7.3.

[5] Notice that the preferences pair $(\succcurlyeq^1, \succcurlyeq^2)$ does play a role in defining aligned acts. Importantly, aligned acts are not necessarily symmetric in the sense that it could be that f and g are $(\succcurlyeq^1, \succcurlyeq^2)$-aligned but g and f are not $(\succcurlyeq^1, \succcurlyeq^2)$-aligned. But, aligned acts are transitive in the sense that if there are n DMs with

This change of perspective also suggests measuring the desirability of acts through probabilities. Specifically, for any act $f \in \mathcal{F}_{a,b}$ with $a \succ^i b$ and any $\alpha \in [0,1]$, we can measure how desirable f is to DM i by asking the DM to express a preference between f and the lottery $(a, \alpha; b, 1-\alpha)$, where $\alpha = 1$ stands for the highest desirability and $\alpha = 0$ stands for the lowest. Combining these two insights, we are led to the following formulation of comparative ambiguity attitudes.

Definition 2. (Prize-Independence Formulation) Fix *any* four prizes $a \succ^1 b$ and $c \succ^2 d$. DM 1 is more ambiguity averse than DM 2 if for any $\alpha \in [0,1]$,

$$f \succsim^1 (a, \alpha; b, 1-\alpha) \quad \Longrightarrow \quad g \succsim^2 (c, \alpha; d, 1-\alpha) \tag{2.1}$$

for all aligned $f \in \mathcal{F}_{a,b}$ and $g \in \mathcal{F}_{c,d}$.

Implicit in Definition 2 is the property that the conclusion of comparison is independent of the four prizes used in aligned acts. Depending on whether this property is satisfied or not, we classify models of decision under uncertainty as either *prize independent* or *prize dependent*, a terminology borrowed from Baillon, Bleichrodt, Li, and Wakker (2019). For prize dependent models, we will still adopt the above formulation but restrict it to certain four prizes (see Definition 5). A detailed discussion of this issue and the models are delayed to section 2.1 and section 2.2.

Definition 2 is stated for an easy comparison with classical definitions, where ambiguous acts are usually contrasted against ambiguity-free acts. However, for applications and for most formal models, an equivalent yet much simpler formulation is available.[6] To state it, first notice that for DM $i \in \{1,2\}$, any $a \succ^i b$ and any act $f \in \mathcal{F}_{a,b}$, there exists under most models a unique number $mp^i(f) \in [0,1]$, named the *matching probability* of f, such that

$$f \quad \sim^i \quad (a, mp^i(f); b, 1 - mp^i(f)).$$

The matching probability of an act $f \in \mathcal{F}_{a,b}$ reflects how a decision maker reduces an ambiguous profile of winning probabilities to an ambiguity-free one (i.e., risk). Using matching probabilities, we expect condition (2.3) to be equivalent to

$$mp^1(f) \leq mp^2(g), \ \textit{for all aligned } f \in \mathcal{F}_{a,b} \textit{ and } g \in \mathcal{F}_{c,d}. \tag{2.2}$$

preferences $\succsim^i$, $i \in \{1, 2, ..., n\}$, and if f and g are $(\succsim^i, \succsim^j)$-aligned and g and h are $(\succsim^j, \succsim^k)$-aligned, then f and h are $(\succsim^i, \succsim^k)$-aligned.

[6] Formally, given preferences $\succsim^1$ and $\succsim^2$ being complete and transitive, the simpler condition is valid if the preferences also satisfy the usual continuity condition, monotonicity condition (state-independent utility), and respect first order stochastic dominance over risk. These conditions are met in most models.

This simplification is indeed valid for all the models we will consider except for the incomplete Bewley preferences. Moreover, when the DMs evaluate lotteries using expected utility, matching probabilities *cardinally* measure the desirability of acts. And, in this situation, the condition can be simply understood as saying that the more ambiguity averse DM always gives a weakly lower evaluation to her act in aligned acts.

To further illustrate the idea, let us compare it with Ghirardato and Marinacci's (2002) widely adopted formulation of comparative ambiguity attitudes. According to that definition, DM 1 is more ambiguity averse than DM 2 if for all $f \in \mathcal{F}$ and $x \in \mathcal{X}$,

$$f \succcurlyeq^1 x \implies f \succcurlyeq^2 x. \tag{2.3}$$

When f is restricted to lotteries, this condition boils down to a Pareto condition over lotteries, which is well known to imply that the two DMs share the same vNM utility if they are expected utility maximizers under risk (see, for example, Harsanyi's (1955) utilitarianism theorem). Using this shared preference over lotteries as a common reference system, condition (2.3) indicates that DM 1 always has a weakly lower evaluation of acts. When the two DMs do have the same preference over lotteries, our condition (2.1) is implied by condition (2.3). However, two innovations are introduced into condition (2.1) to free the DMs from the common-preference restriction. First, instead of using lotteries, we use probabilities on the subjectively best prize as the reference system. Second, using aligned acts and allowing the subjectively best and worst prizes to vary across DMs, we generalize the CMMM condition (2.3) on the small domain of $\mathcal{F}_{a,b}$ and $\mathcal{F}_{c,d}$.

In a sense our innovations relative to condition (2.3) also bring it closer to classical comparative risk analyses, which is indeed what motivated the conditions of Epstein (1999) and Ghirardato and Marinacci's (2002). Specifically, suppose for a moment that $\mathcal{M}$ coincides with the set $\mathbb{R}$ of monetary prizes, and recall that DM 1 is more (absolute) risk averse than DM 2 if for all $x \in \mathcal{X}$ and $m \in \mathcal{M}$,

$$x \succcurlyeq^1 m \implies x \succcurlyeq^2 m. \tag{2.4}$$

Condition (2.4) says that DM 2 prefers a lottery over a sure prize whenever so does DM 1, and similarly, condition (2.3) says that DM 2 prefers an act over a lottery whenever so does DM 1. When lottery x of condition (2.4) takes the form of a sure prize, the condition boils down to the trivial fact that both DMs prefer more over less money. However, when act f of condition (2.3) takes the form of a lottery, the condition implies the nontrivial restriction of common risk preference. By contrast, under our definition, when acts f and g of condition (2.1) take the form of lotteries, the implication remains trivial: both DMs like higher probability on the bigger monetary prize.

To execute the comparison empirically, one needs to first collect matching probabilities $\{(\alpha_i^1, \alpha_i^2)\}_{i=1}^n$ for a sample of aligned acts. Then, a statistical test based on paired data may be employed to reach a conclusion as to which DM is more ambiguity averse.[7] To facilitate the elicitation task, one might consider using cognitively simple aligned acts in the sample, such as those in the form of $f = (a, E; b, E^c)$ and $g = (c, E; d, E^c)$ for various events E (see Baillon, Huang, Selim, and Wakker 2018, and Baillon, Bleichrodt, Li, and Wakker 2019).

We now move to characterization results by applying our formulation of comparative ambiguity attitudes to formal models of decision under uncertainty. For each one of these models, we establish a characterization of comparative ambiguity attitudes without the restriction of common risk preference. When such characterizations (with the common-preference restriction) already exist in the literature, our characterizations agree with them on the essential components. We comment that the comparison is not restricted to be between DMs who are ambiguity averse; many models we consider here can accommodate ambiguity aversion, ambiguity seeking, or a combination or the two.

2.1 Prize Independent Models

Prize independent models possess the formal structure that the conclusion of ambiguity attitudes comparison is independent of the prizes involved in the aligned acts: as explicitly indicated by Definition 2, any four prizes $a \succ^1 b$ and $c \succ^2 d$ will lead to the same conclusion. Models of this kind discussed in this paper include maxmin expected utility (MEU), Bewley preferences, Choquet expected utility, generalized α-maxmin expected utility, and, to a certain degree, homothetic preferences. We present here results for MEU and the closely related Bewley preferences. Readers are referred to section 6 for other models.

To streamline the exposition, we adopt the following notation. Given any utility function u over $\mathcal{X}$ and any act f, $u(f)$ is the *utility profile* of act f under utility function u, that is, it is a simple function that maps $s \in \mathcal{S}$ to $u(f(s))$.[8]

According to MEU (Gilboa and Schmeidler 1989), for DM $i \in \{1, 2\}$ and all acts $f, g \in \mathcal{F}$,

$$f \succsim^i g \iff min_{p \in \mathcal{P}^i} E_p u^i(f) \geq min_{p \in \mathcal{P}^i} E_p u^i(g),$$

[7] For example, one might regress α_i^1 to α_i^2 under various regression settings (e.g., with or without intercept, median regression, etc.) and test if the coefficient is equal to 1, or, conduct a test based on the binomial distribution $B(n, 0.5)$ and the statistic $\sum I(\alpha_i^1 > \alpha_i^2)$, where $I(\cdot)$ is the indicator function. For certain collections of aligned acts, one might also compare indices that aggregate the matching probabilities, like those suggested by Baillon, Huang, Selim, and Wakker (2018).

[8] The set $\{u(f) : f \in \mathcal{F}\}$ coincides with the set $B_0(\Sigma, u(\mathcal{X}))$ of simple functions measurable to Σ and with range in $u(\mathcal{X})$.

where the set $\mathcal{P}^i$ of priors is a compact convex subset of $\Delta(\mathcal{S})$ which is unique, and u^i is a nonconstant affine vNM utility function over $\mathcal{X}$ that is unique up to positive affine transforms. When $\mathcal{P}^i$ is a singleton, MEU reduces to the subjective expected utility (SEU) model. The size of $\mathcal{P}^i$ reflects how much ambiguity DM i takes into account, or, put another way, it reflects to what extent DM i is ambiguity averse. Therefore, the following characterization is a natural one.

Theorem 3. *Given maxmin preferences $\{(u^i, \mathcal{P}^i)\}_{i=1,2}$, the following two statements are equivalent.*

1. *DM 1 is more ambiguity averse than DM 2.*
2. *$\mathcal{P}^1 \supset \mathcal{P}^2$.*

The characterization $\mathcal{P}^1 \supset \mathcal{P}^2$ is widely accepted as the formal content of comparative ambiguity attitudes under MEU. The improvement here is that, while our result puts no restriction on the DMs' risk preferences, most existing results also have $u^1 = u^2$ as part of the statement.

For Bewley preferences (Bewley 2002), the key components are still a vNM utility function u^i and a set of priors $\mathcal{P}^i$, just like in MEU.[9] The Bewley preference $\succcurlyeq^i$ is allowed to be *incomplete* and for all acts $f, g \in \mathcal{F}$,

$$f \succcurlyeq^i g \iff E_p u^i(f) \geq E_p u^i(g), \forall p \in \mathcal{P}^i.$$

The corresponding characterization is stated as follows.

Corollary 4. *Given Bewley preferences $\{(u^i, \mathcal{P}^i)\}_{i=1,2}$, the following two statements are equivalent.*

1. *DM 1 is more ambiguity averse than DM 2.*
2. *$\mathcal{P}^1 \supset \mathcal{P}^2$.*

2.2 Prize Dependent Models

Prize dependent models allow ambiguity attitude to vary with the range of prizes under consideration. It can be natural for behaviors to display such a feature: an investor might become less cautious evaluating an investment opportunity if all the possible final outcomes get improved (by the same amount), because then even the worst case scenario might become profitable. And, it is indeed desirable for models to allow such a feature if we want to study how people's ambiguity attitudes change as their wealth accumulate (see section 4). For

[9]Indeed, the two models can be seen as two standards of rationality that a DM with a set of priors might choose to adhere to (Gilboa, Maccheroni, Marinacci, and Schmeidler 2010).

these models, we need a technical assumption that allows us to study all aspects of a DM's ambiguity attitude using only two prizes.

Assumption B (Boundedness): There are prizes $\overline{m}_1, \underline{m}_1, \overline{m}_2, \underline{m}_2 \in \mathcal{M}$ such that $\overline{m}_1 \succcurlyeq^1 m \succcurlyeq^1 \underline{m}_1$ and $\overline{m}_2 \succcurlyeq^2 m \succcurlyeq^2 \underline{m}_2$ for all $m \in \mathcal{M}$.

This assumption is automatically satisfied when $\mathcal{M}$ consists of finitely many prizes. And, in general, we might still choose *any* $a \succ^1 b$ and $c \succ^2 d$, as in the case of prize independent models, and our characterization results will hold for $\succcurlyeq^1$ and $\succcurlyeq^2$ *restricted* on the sets of acts involving prizes only in the corresponding ranges, which are, $\{f : f(s) \in \Delta\{m : a \succcurlyeq^1 m \succcurlyeq^1 b\}, \forall s \in \mathcal{S}\}$ and $\{f : f(s) \in \Delta\{m : c \succcurlyeq^2 m \succcurlyeq^2 d\}, \forall s \in \mathcal{S}\}$, respectively. Since the reason behind this assumption is technical, we defer a detailed discussion about it and possible relaxations to section 7.2.

For prize dependent models, we will use the following formulation of comparative ambiguity attitudes.

Definition 5. (Prize-Dependent Formulation) Under Assumption B, DM 1 is more ambiguity averse* than DM 2 if for any $\alpha \in [0, 1]$,

$$f \succcurlyeq^1 (\overline{m}_1, \alpha; \underline{m}_1, 1-\alpha) \quad \Longrightarrow \quad g \succcurlyeq^2 (\overline{m}_2, \alpha; \underline{m}_2, 1-\alpha), \tag{2.5}$$

for all aligned $f \in \mathcal{F}_{\overline{m}_1, \underline{m}_1}$ and $g \in \mathcal{F}_{\overline{m}_2, \underline{m}_2}$.

Comparing with Definition 2, Definition 5 requires that only the bounding prizes may appear in the aligned acts used for comparisons. Definition 5 will be used for the smooth ambiguity model and uncertainty averse preferences, the latter including variational preferences and multiplier preferences. We present here the characterization result for the smooth ambiguity model. Readers are referred to section 6 for other models.

Klibanoff, Marinacci, and Mukerji (2005, hereafter KMM) axiomatize the smooth ambiguity model, according to which acts are evaluated as follows[10]

$$f \succcurlyeq^i g \iff \int_\Delta \phi^i\big(E_p u^i(f)\big) d\mu^i(p) \geq \int_\Delta \phi^i\big(E_p u^i(g)\big) d\mu^i(p),$$

where u^i is a nonconstant affine vNM utility function over $\mathcal{X}$, function $\phi^i : u^i(\mathcal{X}) \to \mathbb{R}$ is continuous and strictly increasing, and $\mu^i : \Delta \to [0, 1]$, called *second-order belief*, is a Borel probability distribution over the set of *first-order probabilities* Δ with the weak* topology.

[10]For convenience, the KMM representation is stated here in the Anscombe and Aumann setting. This is consistent with the original axiomatization because KMM do have objective probabilities in their framework. But, the uniqueness property of μ^i does rely on the richer preference structure (that is, over second order acts) assumed in KMM.

When μ^i is degenerate (i.e., it assigns probability 1 to a single $p \in \Delta$), we take ϕ^i to be the identify function. If μ^i is nondegenerate and another triplet $(\hat{u}^i, \hat{\phi}^i, \hat{\mu}^i)$ represent the same preference, it must be that $\hat{\mu}^i = \mu^i$, $\hat{u}^i = k_1 u^i + k_2$ for some $k_1 > 0$ and $k_2 \in \mathbb{R}$, and $\hat{\phi}^i(t) = k_3 \phi^i(\frac{t}{k_1} - \frac{k_2}{k_1}) + k_4$ for some $k_3 > 0$ and $k_4 \in \mathbb{R}$ and all $t \in \hat{u}^i(\mathcal{X})$.

The smooth ambiguity model deviates from SEU because of the transformation function ϕ^i. As ambiguity is captured by a (nondegenerate) second-order belief, a concave ϕ^i indicates that the DM dislikes ambiguity, suggesting an ambiguity averse attitude. The characterization of comparative ambiguity attitudes will thus correspond to a comparison between the concavity of ϕ^1 and ϕ^2. However, to state this result, we need to introduce a technical condition. Viewing first-order probabilities as a subset of the topological vector space of finitely additive signed measures on $(\mathcal{S}, \Sigma)$ with bounded variation, equipped with the weak* topology generated by $B_0(\Sigma, \mathbb{R})$, it makes sense to say that a collection $\{q_n\}_{n=1}^N$ of first-order probabilities are *linearly independent*. Namely, for any $\{\alpha_n\}_{n=1}^N \subset \mathbb{R}$, $\sum_{n=1}^N \alpha_n q_n = 0$ implies $\alpha_n = 0$ for all n. Notice that N can be at most as big as $|\mathcal{S}|$. The technical condition can then be stated as follows.

Assumption CS (Common Support): There is a finite set of linearly independent first-order probabilities $\{q_n\}_{n=1}^N$ such that the support of μ^i is contained in $\{q_n\}_{n=1}^N$, $i = 1, 2$.

This condition can easily be met in a lab where the experimenter controls the extent of ambiguity, or in a conference room where the attendees consider a common set of projections or scenarios that are modeled as first-order probabilities (see Alon and Gayer 2016). In general, it might hold in situations where the DMs share the same information regarding ambiguity. We defer a discussion of this technical assumption after the characterization result, stated next.

Theorem 6. *Given smooth ambiguity models $\{(\phi^i, u^i, \mu^i)\}_{i=1,2}$ with the normalization that $u^1(\overline{m}_1) = u^2(\overline{m}_2)$ and $u^1(\underline{m}_1) = u^2(\underline{m}_2)$, under Assumption B, Assumption CS, and the existence of an interval on which ϕ^1 and ϕ^2 are continuously differentiable, the following two statements are equivalent.*

1. *DM 1 is more ambiguity averse* than DM 2.*
2. *$\phi^1 = \psi \circ \phi^2$ for some concave function $\psi : \phi^2(u^2(\mathcal{X})) \to \mathbb{R}$, and, $\mu^1 = \mu^2$.*

As expected, comparative ambiguity attitudes is reflected through the relative concavity of ϕ^1 to ϕ^2. Somewhat unexpected is the implication $\mu^1 = \mu^2$, though it is actually a natural requirement along the line of separating ambiguity and ambiguity attitudes. Indeed, a more concave ϕ^1 can be in principle compensated by a less "dispersed" μ^1, making DM 1 still willing to give higher evaluations of acts and therefore less ambiguity averse. This possibility is illustrated by an example in the appendix (see section 8.2.1) and it is ruled out here by Assumption CS. KMM's original comparative ambiguity result (see KMM Theorem 2) directly assumes that $\mu^1 = \mu^2$ and this common second-order belief is distributed over two first-order beliefs $\{p, q\} \subset \Delta$ whose supports are disjoint to each other (that is, p and q are *mutually singular*).[11] In section 8.2.2 we also give a proof of the result with $\mu^1 = \mu^2$ being directly assumed without any extra restriction. To summarize, we replace the KMM's assumption of common belief by Assumption CS, which considerably expands the applicability of the comparison, and we show that a common second-order belief is implied by the behavioral condition of comparative ambiguity attitudes.

We remark that Assumption CS can be dropped without affecting the result if preferences can be enriched as in the KMM axiomatization. Specifically, if preferences are extended to the set of all *second-order acts*, which are functions that map first order probabilities Δ to lotteries $\mathcal{X}$, then Theorem 6 holds without Assumption CS (see section 8.2.2). Since our preferences are defined over Anscombe Aumann acts only, which roughly give rise to second order acts that are *affine*, the behavioral condition only implies a common "mean first-order belief", that is, $\int_\Delta p(E)d\mu^1(p) = \int_\Delta p(E)d\mu^2(p)$ for all $E \in \Sigma$. For this reason, Assumption CS is invoked to deliver $\mu^1 = \mu^2$.

If one wants to characterize the relative concavity of ϕ^1 to ϕ^2 without the requirement $\mu^1 = \mu^2$, it can be done in some situations using the method suggested by Baillon, Driesen, and Wakker (2012). Indeed, incorporating their idea into ours will enable both taste and belief to vary across the DMs if we know μ^i is supported on two mutually singular first-order beliefs $\{p^i, q^i\} \subset \Delta$, $i = 1, 2$. However, since this approach does not have the flavor of contrasting ambiguous acts against ambiguity-free acts, we choose to elaborate on it in the appendix (see section 8.2.3).

3 A Decomposition of Certainty Equivalent

For this section, we restrict attention to DM 1 with preference $\succcurlyeq^1$ and DM 2 with preference $\succcurlyeq^2$ whose ambiguity attitudes are comparable in some range of monetary prizes $\mathcal{M} = [m, M]$. Recall that the certainty equivalent (or willingness to pay) of an act $f \in \mathcal{F}$ to DM $i \in \{1, 2\}$ is a sure prize $ce^i(f) \in \mathcal{M}$ such that $f \sim^i ce^i(f)$. Certainty equivalent is expected to exist in reality due to normative reasons, and it does exist for all the models we have considered except for the incomplete Bewley preferences. A natural question is can we decompose the difference between $ce^1(f)$ and $ce^2(f)$ into two parts, one attributable to the DMs' difference in ambiguity attitude, and the other, attributable to their difference in risk attitude?

[11] The KMM assumption cited above is taken from their proof. The actual assumption stated in their Theorem 2 is a bit stronger.

For theoretical works, one can always obtain such a decomposition under formal models, because the certainty equivalent of an act f can be calculated for any hybrid of ambiguity attitude and risk attitude of the two DMs', as our previous results cleared the way for treating ambiguity attitude and risk attitude separately. Even better, for empirical works in the lab or in the field, our previous results also point to an easily implementable *behavioral* decomposition that is free from structural or parametric assumptions. In what follows we elaborate on this behavioral decomposition, the key idea of which lies in the message that matching probability captures ambiguity attitude.

To start with, it is useful to view the evaluation of any act $f \in \mathcal{F}_{M,m}$ with an auxiliary middle step that only reflects ambiguity attitude. Specifically, consider

$$f \underbrace{\sim^i}_{ambiguity} (M, mp^i(f); m, 1 - mp^i(f)) \underbrace{\sim^i}_{risk} ce^i(f). \tag{3.1}$$

That is, the first indifference relation reduces ambiguity to risk, which reflects ambiguity attitude, and the second indifference relation further reduces risk to certainty, which reflects risk attitude. To get a decomposition of $ce^2(f) - ce^1(f)$, we need four certainty equivalents to be collected in two steps. In step 1, we incentivize the two DMs to find $mp^i(f)$ and $ce^i(f)$, and, in step 2, we incentivize them again to find the certainty equivalents but now using the *exchanged* matching probability, namely, $ce^1\big((M, mp^2(f); m, 1 - mp^2(f))\big)$ and $ce^2\big((M, mp^1(f); m, 1 - mp^1(f))\big)$. Notice that when DM 1 and DM 2 exchanges matching probabilities, it is as if each of them took the ambiguity attitude of the opponent's but still kept her own risk attitude. Therefore we can interpret the four certainty equivalents as shown in Table 1.

	DM 1 Ambiguity	DM 2 Ambiguity
DM 1 Risk	$ce^1(f)$	$ce^1\big((M, mp^2(f); m, 1 - mp^2(f))\big)$
DM 2 Risk	$ce^2\big((M, mp^1(f); m, 1 - mp^1(f))\big)$	$ce^2(f)$

Table 1: Decomposition for act $f \in \mathcal{F}_{M,m}$

Notes: Here we show the four certainty equivalents of act f according to each possible combination of risk attitude and ambiguity attitude. The two off-diagonal entries are obtained by incentivizing the DMs to give monetary evaluations based on their opponent's matching probability.

According to Table 1, the discrepancies $ce^1\big((M, mp^2(f); m, 1 - mp^2(f))\big) - ce^1(f)$ and

$ce^2(f) - ce^2\big((M, mp^1(f); m, 1 - mp^1(f))\big)$ should be attributed to the DMs' difference in ambiguity attitude, while the discrepancies $ce^2(f) - ce^1\big((M, mp^2(f); m, 1 - mp^2(f))\big)$ and $ce^2\big((M, mp^1(f); m, 1-mp^1(f))\big) - ce^1(f)$, to the DM's difference in risk attitude. Notice that the corresponding parts are 0 when the DMs do share the same risk attitude or ambiguity attitude. To decompose $ce^2(f) - ce^1(f)$, we might take simple averages and conclude that

$$\begin{aligned} & \tfrac{1}{2}\left(ce^2\big((M, mp^1(f); m, 1 - mp^1(f))\big) - ce^1(f)\right) \\ + \; & \tfrac{1}{2}\left(ce^2(f) - ce^1\big((M, mp^2(f); m, 1 - mp^2(f))\big)\right) \end{aligned} \tag{3.2}$$

is attributed to the difference in risk attitudes and

$$\begin{aligned} & \tfrac{1}{2}\left(ce^1\big((M, mp^2(f); m, 1 - mp^2(f))\big) - ce^1(f)\right) \\ + \; & \tfrac{1}{2}\left(ce^2(f) - ce^2\big((M, mp^1(f); m, 1 - mp^1(f))\big)\right) \end{aligned} \tag{3.3}$$

is attributed to the difference in ambiguity attitudes.

For general act $f \in \mathcal{F}$, a behavioral decomposition is also available, albeit a bit more complicated. We start with the observation made in Wang (2019) that for each state $s \in \mathcal{S}$, DM i is indifferent between the lottery $f(s)$ and a unique binary lottery $f^i_{M,m}(s)$ in $\Delta(\{M, m\})$, and, by replacing $f(s)$ with $f^i_{M,m}(s)$ state by state we get a unique *reduced act* $f^i_{M,m} \in \mathcal{F}_{M,m}$ such that $f \sim^i f^i_{M,m}$ and $f^i_{M,m}(s) \sim^i f(s)$ for all $s \in S$, $i \in \{1, 2\}$.[12] This state-by-state replacement procedure is impacted only by the DMs' risk attitudes and it is redundant if f is already in $\mathcal{F}_{M,m}$. Once we have $f^i_{M,m}$, the existing two-step evaluation (3.1) can proceed. Therefore, this observation amounts to prefixing another auxiliary step to the existing procedure, illustrated as follows.

$$f \underbrace{\sim^i}_{risk} f^i_{M,m} \underbrace{\sim^i}_{ambiguity} (M, mp^i(f^i_{M,m}); m, 1 - mp^i(f^i_{M,m})) \underbrace{\sim^i}_{risk} ce^i(f). \tag{3.4}$$

Due to this extra initial indifference relation, we need to adjust the method in collecting the four certainty equivalents. Specifically, in step 1, we elicit $f^i_{M,m}$ and $ce^i(f)$, $i \in \{1, 2\}$. In step 2, the DMs *switch* $f^i_{M,m}$ and give matching probabilities accordingly, from which $mp^1(f^2_{M,m})$ and $mp^2(f^1_{M,m})$ are obtained. Finally, in step 3, the DMs *switch* matching probabilities and re-evaluate the certainty equivalents accordingly, from which $ce^1\big((M, mp^2(f^1_{M,m}); m, 1 - mp^2(f^1_{M,m}))\big)$ and $ce^2\big((M, mp^1(f^2_{M,m}); m, 1 - mp^1(f^2_{M,m}))\big)$ are obtained. Notice that in step 3 it is as if the DMs switch ambiguity attitude but still maintain their own risk attitudes. Therefore we can interpret the four certainty equivalents as shown in Table 2.

[12] This procedure is valid for all the models we considered except for the incomplete Bewley preferences.

	DM 1 Ambiguity	DM 2 Ambiguity
DM 1 Risk	$ce^1(f)$	$ce^1\left(\begin{pmatrix} M, mp^2(f^1_{M,m}); \\ m, 1-mp^2(f^1_{M,m}) \end{pmatrix}\right)$
DM 2 Risk	$ce^2\left(\begin{pmatrix} M, mp^1(f^2_{M,m}); \\ m, 1-mp^1(f^2_{M,m}) \end{pmatrix}\right)$	$ce^2(f)$

Table 2: Decomposition for generic act $f \in \mathcal{F}$

Notes: Here we show the four certainty equivalents of act f according to each possible combination of risk attitude and ambiguity attitude. The two off-diagonal entries are monetary evaluation(s) of DM $i \in \{1,2\}$ based on her opponent's matching probability of DM i's own reduced act $f^i_{M,m} \in \mathcal{F}_{M,m}$, which satisfies $f \sim^i f^i_{M,m}$ and $f^i_{M,m}(s) \sim^i f(s)$ for all $s \in \mathcal{S}$.

According to Table 2, the discrepancies $ce^1\big((M, mp^2(f^1_{M,m}); m, 1-mp^2(f^1_{M,m}))\big) - ce^1(f)$ and $ce^2(f) - ce^2\big((M, mp^1(f^2_{M,m}); m, 1-mp^1(f^2_{M,m}))\big)$ should be attributed to the difference in ambiguity attitude, while the discrepancies $ce^2(f) - ce^1\big((M, mp^2(f^1_{M,m}); m, 1-mp^2(f^1_{M,m}))\big)$ and $ce^2\big((M, mp^1(f^2_{M,m}); m, 1-mp^1(f^2_{M,m}))\big) - ce^1(f)$, to the difference in risk attitude. Notice that the corresponding parts are indeed 0 when the DMs share the same risk attitude or ambiguity attitude. To decompose $ce^2(f) - ce^1(f)$, we might take simple averages and conclude that

$$\begin{aligned} & \tfrac{1}{2}\left(ce^2\big((M, mp^1(f^2_{M,m}); m, 1-mp^1(f^2_{M,m}))\big) - ce^1(f)\right) \\ + \quad & \tfrac{1}{2}\left(ce^2(f) - ce^1\big((M, mp^2(f^1_{M,m}); m, 1-mp^2(f^1_{M,m}))\big)\right) \end{aligned} \tag{3.5}$$

is attributed to the difference in risk attitudes and

$$\begin{aligned} & \tfrac{1}{2}\left(ce^1\big((M, mp^2(f^1_{M,m}); m, 1-mp^2(f^1_{M,m}))\big) - ce^1(f)\right) \\ + \quad & \tfrac{1}{2}\left(ce^2(f) - ce^2\big((M, mp^1(f^2_{M,m}); m, 1-mp^1(f^2_{M,m}))\big)\right) \end{aligned} \tag{3.6}$$

is attributed to the difference in ambiguity attitudes.

For a simple application, let us return to our motivating example of a trade between a venture capitalist and an entrepreneur.

Example 7. An entrepreneur with preference $\succcurlyeq^1$ has a start-up project, modeled as an ambiguous act f that yields monetary outcome M in a success state and m in a failure state. A venture capitalist, with preference $\succcurlyeq^2$, seeks to acquire (maybe part of) the ownership of the project. Leading to this deal, two channels are at work. On one hand, since the entrepreneur knows the firm better and tends to be optimistic (that is, less ambiguity averse), she gives an estimate $mp^1(f)$ of the chance of success much higher than the assessment $mp^2(f)$ made by the venture capitalist. On the other hand, since the venture capitalist usually has a diversified portfolio, she is much more risk tolerant, leading to a final evaluation $ce^2(f)$ that is nevertheless higher than the entrepreneur's evaluation $ce^1(f)$. The two channels are summarized as follows.

$$\begin{array}{rcccccc} \textit{entrepreneur}: & f & \overbrace{\sim^1}^{\textit{ambiguity}} & (M, mp^1(f); m, 1-mp^1(f)) & \overbrace{\sim^1}^{\textit{risk}} & ce^1(f) \\ \textit{venture capitalist}: & f & \sim^2 & (M, mp^2(f); m, 1-mp^2(f)) & \sim^2 & ce^2(f) \end{array}$$

To disentangle the impact of the two channels, we ask the two of them to *switch* their probabilistic assessments and give new monetary evaluations accordingly, which are denoted by $ce^1((M, mp^2(f); m, 1-mp^2(f)))$ and $ce^2\big((M, mp^1(f); m, 1-mp^1(f))\big)$, respectively. Then, the overall premium $ce^2(f)-ce^1(f)$ can be decomposed into a negative part, shown in formula (3.3), due to the venture capitalist being more ambiguity averse, and a positive part, shown in formula (3.2), due to the venture capitalist being less risk averse.

4 Absolute and Relative Ambiguity Attitudes

How to model a DM's preference if attitude towards uncertainty changes with wealth level? Starting from Arrow (1971), this question has been investigated for risk attitude in a series of papers, resulting in widely used functional forms like CARA and CRRA. Empirical evidences also suggest that ambiguity attitude may change as the underlying wealth level varies (see Baillon and Placido 2019). A small literature is indeed dedicated to pinning down the behavioral content and formal characterizations of absolute and relative ambiguity attitudes. However, all existing results seem to have certain constraints. Roughly speaking, one strand of this literature does not disentangle risk and ambiguity attitudes (see Cherbonnier and Gollier 2015), one strand leverages a notion of constant mixtures to work more or less with shifts of utility profile instead of shifts of wealth (see Grant and Polak 2013 and Xue 2018), while another strand, though working directly with wealth level, imposes constant absolute risk attitude or constant relative risk attitude on the DM (see Cerreia-Vioglio, Maccheroni, and Marinacci 2019, hereafter CMM). Readers are referred to CMM for a detailed literature review. The approach we take to this question is closely related to CMM's, and our conceptual novelty lies in applying our formulation of comparative ambiguity attitudes to a single DM, who is to be labeled by various wealth levels. The resulting solution is arguably more general and it is without the aforementioned constrains.

In this section we focus on monetary prizes, that is, $\mathcal{M} = \mathbb{R}$ or $\mathcal{M} = \mathbb{R}_{++}$. For any act $f \in \mathcal{F}$ and any wealth level $w \in \mathbb{R}$, f^{+w} denotes act f with all prizes in it being shifted by w, that is, $f^{+w}(s)(c) = f(s)(c-w)$ for all $s \in S$ and $c \in \mathbb{R}$. And for any act $f \in \mathcal{F}$ and any wealth level $w > 0$, $f^{\times w}$ denotes act f with all prizes in it being multiplied by w, that is, $f^{\times w}(s)(c) = f(s)(c/w)$ for all $s \in S$ and $c \in \mathbb{R}$. Notice that for any prize $a > 0$ and any act $f \in \mathcal{F}_{a,0}$, acts f^{+w} and f^{+v} are aligned for all wealth levels $w > v$. Similarly, for any number $k > 1$ and any act $f \in \mathcal{F}_{k,1}$, acts $f^{\times w}$ and $f^{\times v}$ are aligned for all $w > v > 0$.

Throughout the section $\succcurlyeq$ is a *rational preference* as described by Theorem 1 of CMM, which encompasses all the models we considered so far except for Bewley preferences. In effect, $\succcurlyeq$ can be represented by a pair (u, I) such that $f \succcurlyeq g \Leftrightarrow I(u(f)) \geq I(u(g))$, for which $u : X \to \mathbb{R}$ is a nonconstant and affine vNM utility function, with $u(f)$ standing for the utility profile of any act f, and $I : B_0(\Sigma, u(X)) \to \mathbb{R}$ is a normalized, monotone, and continuous function. The terminologies are defined as follows. For any constant $c \in \mathbb{R}$, the bold $\mathbf{c}$ denotes the constant function that maps all state to c. We say $I(\cdot)$ is *normalized* if $I(\mathbf{c}) = c$ for any $c \in u(X)$, *monotone* if $I(\psi) \geq I(\phi)$ *for all* $\psi, \phi \in B_0(\Sigma, u(X))$ *and* $\psi \geq \phi$, *constant additive* if $I(\mathbf{c} + \psi) = c + I(\psi)$ for any $\psi \in B_0(\Sigma, u(X))$ and constant function $\mathbf{c}$ such that $\mathbf{c} + \psi \in B_0(\Sigma, u(X))$, and *superhomogeneous* (respectively, *subhomogeneous*) if $I(\lambda\psi) \geq \lambda I(\psi)$ (respectively, $I(\lambda\psi) \leq \lambda I(\psi)$) for all $\psi \in B_0(\Sigma, u(X))$ and $\lambda \in (0,1)$ such that $\lambda\psi \in B_0(\Sigma, u(X))$.

Definition 8. 1. Preference $\succcurlyeq$ satisfies *decreasing absolute ambiguity aversion (DAAA)* if for any two wealth levels $w > v$, any prize $a > 0$, and any probability $p \in [0,1]$,

$$f^{+v} \succcurlyeq (v + a, p; v, 1-p) \implies f^{+w} \succcurlyeq (w + a, p; w, 1-p), \tag{4.1}$$

for all $f \in \mathcal{F}_{a,0}$.

2. Preference $\succcurlyeq$ satisfies *decreasing relative ambiguity aversion (DRAA)* if for any two wealth levels $w > v > 0$, any number $k > 1$, and any probability $p \in [0,1]$,

$$f^{\times v} \succcurlyeq (kv, p; v, 1-p) \implies f^{\times w} \succcurlyeq (kw, p; w, 1-p), \tag{4.2}$$

for all $f \in \mathcal{F}_{k,1}$.

Hence, the formulation of DAAA and DRAA mimics that of comparative ambiguity attitudes by treating the same DM at different wealth levels as distinct DMs. Increasing absolute ambiguity aversion (IAAA) and increasing relative ambiguity aversion (IRAA) are defined by replacing the implies symbol $\implies$ of condition (4.1) and condition (4.2) with the implied-by symbol $\Longleftarrow$. Constant absolute (relative) ambiguity aversion holds if both DAAA and IAAA (DRAA and IRAA) hold.

Comparing with CMM's formulation, the difference is that while our condition is based on Definition 2 and therefore involves only binary lotteries, CMM use Ghirardato and Marinacci's (2002) classic condition (2.3) and have all lotteries involved. In this sense, our formulation restricts CMM's to a smaller domain. And, in doing this, our formulation essentially frees utility function $u(\cdot)$ from any restriction, while CMM's implicitly imposes CARA on $u(\cdot)$ in the case of DAAA, and CRRA on $u(\cdot)$ in the case of DRAA. The mechanism that allows us to achieve this separation remains the same as in the case of comparative ambiguity attitudes. In short, when the act f of Definition 8 is ambiguity-free and henceforth a binary lottery, condition (4.1) and condition (4.2) trivially hold. Therefore, it leads to no extra restriction.

The interaction between ambiguity aversion and wealth level can also be studied through matching probabilities. Indeed, the DAAA condition (4.1) is equivalent to requiring matching probabilities to increase as wealth increases, that is, for all $w > v$ and $a > 0$, $mp(f^{+v}) \leq mp(f^{+w})$ for all $f \in \mathcal{F}_{a,0}$. In comparison, CMM's formulation of DAAA amounts to requiring certainty equivalents to increase as wealth increases, which necessarily carries the information of both ambiguity attitude and risk attitude. Once again, matching probability manifests as a concept that captures ambiguity attitude independent of risk attitude.

For absolute ambiguity attitude, we state a characterization result of DAAA. One only needs to switch the corresponding inequality signs for obtain a result for IAAA.

Theorem 9. *Given rational preferences $\succcurlyeq^1$ and $\succcurlyeq^2$, the following statements are equivalent.*

1. Decreasing absolute ambiguity aversion holds.

2. (Increasing matching probability) For any two wealth levels $w > v$ and any prize $a > 0$, $mp(f^{+w}) \geq mp(f^{+v})$ for all $f \in \mathcal{F}_{a,0}$.

3. For any two wealth levels $w > v$, any prize $a > 0$, and any $\psi \in B_0(\Sigma, [0, u(v + a) - u(v)])$, we have

$$I(u(w) + \lambda\psi) - u(w) \geq \lambda\big(I(u(v) + \psi) - u(v)\big),$$

where $\lambda = [u(w + a) - u(w)]/[u(v + a) - u(v)]$.

For relative ambiguity aversion, we state a result for DRAA. Again, one obtains a result for IRAA by switching the corresponding inequality signs.

Theorem 10. *Given rational preferences $\succcurlyeq^1$ and $\succcurlyeq^2$, the following statements are equivalent.*

1. Decreasing relative ambiguity aversion holds.

2. (Increasing matching probability) For any two wealth levels $w > v > 0$ and any $k > 1$, $mp(f^{\times w}) \geq mp(f^{\times v})$ for all $f \in \mathcal{F}_{k,1}$.

3. For any two wealth levels $w > v > 0$, any $k > 1$, and any $\psi \in B_0(\Sigma, [0, u(kv) - u(v)])$, we have

$$I(u(w) + \lambda\psi) - u(w) \geq \lambda\big(I(u(v) + \psi) - u(v)\big),$$

where $\lambda = [u(kw) - u(w)]/[u(kv) - u(v)]$.

To answer the classic question, these results provide guidance as to how to model the interaction between ambiguity attitude and wealth level. For example, if we set $I(\cdot)$ to be constant additive, then $\succcurlyeq$ is DAAA if $I(\cdot)$ is superhomogeneous and $\succcurlyeq$ is risk averse (so $\lambda \leq 1$) or if $I(\cdot)$ is subhomogeneous and $\succcurlyeq$ is risk loving ($\lambda \geq 1$), and, symmetrically, $\succcurlyeq$ is IAAA if $I(\cdot)$ is superhomogeneous and $\succcurlyeq$ is risk loving ($\lambda \geq 1$) or if $I(\cdot)$ is subhomogeneous and $\succcurlyeq$ is risk averse ($\lambda \leq 1$). This guidance is also more general than provided by CMM. Indeed, when $\succcurlyeq$ is CARA in the risk domain, Theorem 9 implies Theorem 2 of CMM, and, when $\succcurlyeq$ is CRRA in the risk domain, Theorem 10 implies Theorem 4 of CMM.

Due to Baillon and Placido's (2019) evidences supporting DAAA, we give some examples of DAAA with no restriction on risk attitude. The first example models DAAA (and DRAA) through a sequence of shrinking sets of priors, as in maxmin preferences.

Example. (DAAA and DRAA for maxmin preferences) Take wealth levels $w_1 < w_2 < ... < w_n$ and sets of priors $\mathcal{P}_0 \supset \mathcal{P}_1 \supset ... \supset \mathcal{P}_n$. The interpretation is going to be that the DM's set of priors shrinks as wealth level increases. For any $\psi \in B_0(\Sigma, u(X))$, we define $I(\psi) = min_{p \in \mathcal{P}_n} E_p\psi$ if $min_s\psi(s) \geq u(w_n)$, $I(\psi) = min_{p \in \mathcal{P}_0} E_p\psi$ if $min_s\psi(s) < u(w_1)$, and

$$I(\psi) = \frac{u(w_{m+1}) - min_s\psi(s)}{u(w_{m+1}) - u(w_m)} min_{p \in \mathcal{P}_m} E_p\psi + \frac{min_s\psi(s) - u(w_m)}{u(w_{m+1}) - u(w_m)} min_{p \in \mathcal{P}_{m+1}} E_p\psi$$

if $min_s\psi(s) \in [u(w_m), u(w_{m+1}))$ for some $m \in \{1, 2, ..., n-1\}$. This I is normalized, monotone, and continuous.

The second example models DAAA (and DRAA) through a sequence of increasing weights associated with the optimistic scenario as opposed to the pessimistic one, as in α-maxmin preferences.

Example. (DAAA and DRAA for α-maxmin preferences) Fix a set $\mathcal{P}$ of priors, and take wealth levels $w_1 < w_2 < ... < w_n$ and weights $0 \leq \alpha_0 < \alpha_1 < ... < \alpha_n \leq 1$. The interpretation is going to be that the weight the DM puts on the optimistic scenario increases as wealth level increases. For any $\psi \in B_0(\Sigma, u(X))$, we define

$$I(\psi) = \alpha(\psi) max_{p \in \mathcal{P}} E_p\psi + (1 - \alpha(\psi)) min_{p \in \mathcal{P}} E_p\psi,$$

where $\alpha(\psi) = \alpha_n$ if $min_s\psi(s) \geq u(w_n)$, $\alpha(\psi) = \alpha_0$ if $min_s\psi(s) < u(w_1)$, and

$$\alpha(\psi) = \frac{u(w_{m+1}) - min_s\psi(s)}{u(w_{m+1}) - u(w_m)}\alpha_m + \frac{min_s\psi(s) - u(w_m)}{u(w_{m+1}) - u(w_m)}\alpha_{m+1}$$

if $min_s\psi(s) \in [u(w_m), u(w_{m+1}))$ for some $m \in \{1, 2, ..., n-1\}$. This I is normalized, monotone, and continuous.

The next example addresses absolute ambiguity attitude in the smooth ambiguity model, which turns out to be quite handy.

Example. (DAAA and IAAA for the smooth ambiguity model) Fix a second order belief μ over $\Delta(S)$. Then, for the smooth ambiguity model we have

$$I(\psi) = \phi^{-1}\bigg(\int_\Delta \phi(E_p\psi)d\mu(p)\bigg)$$

for any function $\psi \in B_0(\Sigma, u(X))$. Suppose $\phi(x) = -e^{-rx}$ for some $r > 0$. In this case, $I(\cdot)$ is constant additive and superhomogeneous. Therefore, $\succcurlyeq$ is DAAA under risk aversion ($\lambda \leq 1$ for all $w > v$ and $a > 0$) and IAAA under risk loving ($\lambda \geq 1$ for all $w > v$ and $a > 0$). Suppose $\phi(x) = e^{rx}$ for some $r > 0$. In this case, $I(\cdot)$ is constant additive and subhomogeneous. Therefore, $\succcurlyeq$ is DAAA under risk loving and IAAA under risk aversion.

5 Applications

5.1 Risk Attitude and Ambiguity Attitude May Have Opposite Effects

A clean separation between ambiguity attitude and risk attitude is desirable to a researcher who wants to contrast the economic impact of ambiguity attitude against that of risk attitude. This contrast is especially interesting when ambiguity attitude and risk attitude affect optimal behavior in opposite directions. Here we discuss by example a general mechanism under which such a phenomenon is possible. Roughly, while higher risk aversion usually induces the DM to strive for a safer prospect, higher ambiguity aversion might only make moderate actions seem less effective, pushing the DM to take extreme actions, that is, either making a high effort to reduce uncertainty or making a low effort and just relying on luck. This mechanism might be at work in situations where the DM needs to make an effort to get higher chances of good outcomes (e.g., the moral hazard problem) or lower chances of shocks (e.g., insurance problems). To illustrate, let us consider the following example.

Example. A DM has a project which might fail due to either cause A or cause B. The project brings monetary outcome \$$M$ when succeeds and \$0 when fails. Suppose A and B happen with probability α and β, respectively, and they do not happen concurrently. Further suppose that A is more likely to happen than B (i.e., $\alpha > \beta$) and failure is unlikely (i.e., $\alpha + \beta < 0.5$). It is possible to take preemptive measures to prevent A and B from happening, costing \$$c$ each. The DM decides whether to take preemptive measures against cause A, or cause B, or both, or none.

We consider three situations: the DM is risk neutral and ambiguity neutral, the DM is ambiguity neutral but risk averse (modeled by mean-variance preference with risk aversion coefficient θ), and the DM is risk neutral but ambiguity averse (modeled by MEU with the extra scenario where A and B happen with probability β and α, respectively). Table 3 summarizes the payoffs of taking each action in each situation.

Number of	DM is neutral to risk (R) or ambiguity (A)		
Preventions	R & A Neutral	A Neutral	R Neutral
None	$(1-\alpha-\beta)M$	$(1-\alpha-\beta)M$ $-\frac{1}{2}\theta(\alpha+\beta)(1-\alpha-\beta)M^2$	$(1-\alpha-\beta)M$
One	$(1-\beta)M - c$	$(1-\beta)M - c$ $-\frac{1}{2}\theta\beta(1-\beta)M^2$	$(1-\alpha)M - c$
Both	$M - 2c$	$M - 2c$	$M - 2c$

Table 3: Payoff Matrix

Note: The table summarizes the payoffs of taking each action to the three types of DMs. The risk and ambiguity neutral type evaluates action using expected value. The ambiguity neutral (but risk averse) type uses mean-variance model with risk aversion coefficient θ. The risk neutral (but ambiguity averse) type uses minimum expected value, a special case of maxmin expected utility. Notice that, under no ambiguity, when the DM takes preemptive measures against only one cause, the one that happens with higher probability will be chosen.

Suppose $\alpha M > c$ and $\beta M < c$. Then the risk and ambiguity neutral DM chooses to take the moderate action, that is, to only prevent cause A from happening. When the DM becomes risk averse yet ambiguity neutral, relative to taking no preemptive measures, the moderate action becomes even more desirable due to $\beta(1-\beta) < (\alpha+\beta)(1-\alpha-\beta)$, and so does preventing both causes due to risk aversion. That is, the DM overall has more incentive to reduce uncertainty. However, this is not necessarily the case when the DM becomes ambiguity averse yet risk neutral. Here, the payoff of preventing one cause becomes smaller (from αM to βM), as the DM will always worry that the cause not being prevented is actually the one that happens with higher probability. Since $\beta M < c$, preventing one cause will never be adopted because it is worse than doing nothing. Therefore, the DM will choose to take no preemptive measures at all if $(\alpha+\beta)M < 2c$, and choose to prevent both causes if $(\alpha+\beta)M > 2c$.

5.2 Portfolio Choice under Robust Mean-Variance Preferences

Maccheroni, Marinacci, and Ruffino (2013, MMR) propose the class of robust mean-variance preferences and justify it as the quadratic approximation of the smooth ambiguity model. An asset is modeled as a real-valued element in $L^2(\mathcal{S}, \Sigma, p)$, that is, its square is integrable with respect to a first order probability distribution p. Ambiguity is modeled as a second order probability distribution μ over the set of first order probability distributions that are absolute continuous with respect to p, with Radon-Nikodym derivatives in $L^2(\mathcal{S}, \Sigma, p)$, and and its *reduction* is p, that is, $\int_\Delta q(E)d\mu(q) = p(E)$ for all $E \in \Sigma$. According to robust mean-variance preferences, any asset r is evaluated as

$$E_p r - \frac{1}{2}\theta_r E_\mu\big(\sigma_q^2(r)\big) - \frac{1}{2}\theta_a \sigma_\mu^2(E_q r), \tag{5.1}$$

where $E_p r$ is the overall mean return, $E_\mu\big(\sigma_q^2(r)\big)$ is the mean of variance $\sigma_q^2(r)$ across scenarios q, $\sigma_\mu^2(E_q r)$ is the variance of mean return $E_q r$ across scenarios q, and θ_r and θ_a reflects risk attitude and ambiguity attitude, respectively. And, in approximating the smooth ambiguity model, θ_r and θ_a are the (local) Arrow-Pratt coefficients $-\frac{u''}{u'}$ of the vNM utility and $-\frac{\phi''}{\phi'}$ of the transformation function, respectively. By our characterization result on the smooth ambiguity model (see Theorem 6), we are justified to interpret θ_r and θ_a separately.[13]

We follow MMR to consider portfolio choice problems under robust mean-variance preferences. Suppose there are two types of assets: a risk-free and ambiguity-free asset r_f and an ambiguous asset r. The agent chooses amount $w \in \mathbb{R}$ on the ambiguous asset to maximize the evaluation of portfolio $(1-w)r_f + wr$. MMR show that the optimal $\hat{w}$ satisfies

$$\hat{w} = \frac{E_p r - r_f}{\theta_r E_\mu\big(\sigma_q^2(r)\big) + \theta_a \sigma_\mu^2(E_q r)}.$$

[13] For a direct justification, Definition 2 does not apply since robust mean-variance preferences are not in the Anscombe Aumann framework. However, a direct justification is possible at least in the lab using the following behavioral comparison, which leverages the simple structure of the model. First, let the two DMs know μ which is distributed on two mutually singular first order probabilities q_1 and q_2. Then, take an asset r that yields return $x\%$ on $supp(q_1)$ and $y\%$ on $supp(q_2)$ with $x \neq y$, and ask DM $i \in \{1,2\}$ to find asset r_i such that $r_i \sim^i r$ and it has the same return across all states. Since r bears ambiguity but no risk, henceforth we have $\theta_a^1 \leq \theta_a^2$ if and only if r_1's return is lower than that of r_2.

With risk attitude being separated from ambiguity attitude, a researcher can conclude that higher risk aversion or higher ambiguity aversion or *both* lead to a smaller position $\hat{w}$ on the ambiguous asset.

When there are three types of assets — a risk-free and ambiguity-free asset r_f, a purely risky asset r_m, and an ambiguous asset r_e — a researcher can contrast the impacts of risk aversion θ_r and ambiguity aversion θ_a on the optimal portfolio. The agent chooses fractions $(w_f, w_m, w_e) \in \mathbb{R}^3$ with $w_f + w_m + w_e = 1$ to maximize the evaluation of portfolio $w_f r_f + w_m r_m + w_e r_e$. MMR show that the optimal solution $(\hat{w}_f, \hat{w}_m, \hat{w}_e)$ can be described as follows

$$\hat{w}_m = \frac{AD - BH}{CD - H^2}, \hat{w}_e = \frac{BC - AH}{CD - H^2}, \text{ and } \hat{w}_f = 1 - \hat{w}_m - \hat{w}_e,$$

where $A = E_p(r_m - r_f)$, $B = E_p(r_e - r_f)$, $C = \theta_r \sigma_p^2(r_m)$, $D = \theta_r E_\mu\big(\sigma_q^2(r_e)\big) + \theta_a \sigma_\mu^2(E_q r_e)$, and $H = \theta_r \sigma_p(r_m, r_e)$. Roughly, A and B are respectively the "uncertainty premium" of r_m and r_e, C and D are respectively the "uncertainty penalty" of r_m and r_e, and H measures the impact of the covariance of r_m and r_e.

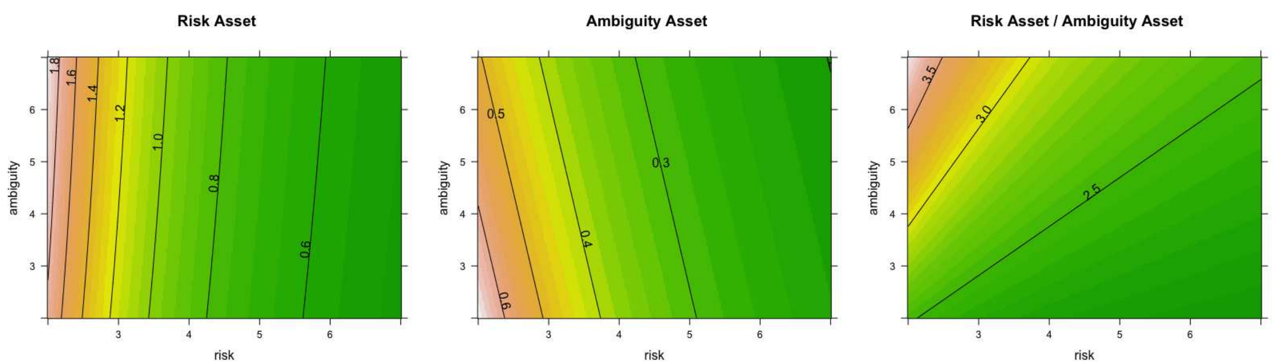

Figure 5.1: Optimal Portfolio

Note: From left to right are optimal risk asset $\hat{w}_m$, optimal ambiguity asset $\hat{w}_e$, and their ratio $\hat{w}_m/\hat{w}_e$. The x-axis and y-axis denote risk aversion parameter θ_r and ambiguity aversion parameter θ_a, respectively. The environment parameters are set at $r_f = 0.05$, $E_p r_m = 0.1$, $E_p r_e = 0.15$, $\sigma_p^2(r_m) = 0.01$, $E_\mu\big(\sigma_q^2(r_e)\big) = 0.04$, $\sigma_\mu^2(E_q r_e) = 0.005$, and $\sigma_p(r_m, r_e) = 0.01061$ with an overall correlation $\rho_p(r_m, r_e)$ between r_m and r_e equal to 0.5.

To illustrate the optimal portfolio, consider the numerical example summarized in Figure 5.1. Here, higher ambiguity aversion θ_a makes ambiguous asset less attractive ($\hat{w}_m$ increases and $\hat{w}_e$ decreases), while higher risk aversion θ_r makes both type of uncertain assets less attractive (both $\hat{w}_m$ and $\hat{w}_e$ decrease) with the impact on risk asset more pronounced ($\hat{w}_m/\hat{w}_e$ decreases). And, risk aversion may offset ambiguity aversion in affecting the optimal balance $\hat{w}_m/\hat{w}_e$ between the risk asset and ambiguous asset, as shown in the right panel.

6 Comparative Ambiguity Attitudes for Other Models

In this section we apply our formulation of comparative ambiguity attitudes to some other popular models of decision under uncertainty. Again, when characterization results (with the common-preference restriction) already exist in the literature, our characterizations agree with them on the essential components.

6.1 Choquet Expected Utility

Schmeidler (1989) axiomatizes the Choquet expected utility (CEU) model, according to which for DM $i \in \{1, 2\}$ and all acts $f, g \in \mathcal{F}$,

$$f \succsim^i g \iff \int_{\mathcal{S}} u^i(f) dv^i \geq \int_{\mathcal{S}} u^i(g) dv^i,$$

where $v^i : \Sigma \to [0, 1]$ is a capacity that is unique, u^i is a nonconstant affine vNM utility function over $\mathcal{X}$ that is unique up to positive affine transforms, and the integral is the Choquet integral.[14] The deviation from SEU is that the prior is replaced by a capacity, which reflects the ambiguity attitude.

Proposition 11. *Given Choquet expected utility models $\{(u^i, v^i)\}_{i=1,2}$, the following three statements are equivalent.*

1. *DM 1 is more ambiguity averse than DM 2.*
2. *Condition (1.2) holds.*
3. *(Epstein 1999) $v^1 \leq v^2$.*

This result also shows that for CEU the simpler behavioral condition (1.2) is sufficient. The reason is roughly that CEU generally has a much lower degree of freedom than, for example, MEU, requiring a less rich condition to discipline its formal components (see also Basu and Echenique 2018). When DM i is uncertainty averse according to Schmeidler's (1989) definition, the capacity v^i has a nonempty core $core(v^i) \equiv \{p \in \Delta : p(E) \geq v^i(E), \forall E \in \Sigma\}$, whose size naturally measures the extent to which DM i is ambiguity averse. If both DMs are uncertainty averse, then $v^1 \leq v^2$ is equivalent to $core(v^1) \supset core(v^2)$.

[14]Function $v : \Sigma \to [0, 1]$ is a capacity if $v(\emptyset) = 0$, $v(\mathcal{S}) = 1$, and $v(E) \leq v(F)$ for all events $E, F \in \Sigma$ and $E \subset F$.

6.2 Generalized α-Maxmin Expected Utility

Ghirardato, Maccheroni, and Marinacci (2004, hereafter GMM) axiomatizes a generalized α-maxmin expected utility, which evaluates acts as follows

$$\mathcal{F} \ni f \mapsto \alpha^i[u^i(f)]min_{p\in\mathcal{P}^i}E_p u^i(f) + (1-\alpha^i[u^i(f)])max_{p\in\mathcal{P}^i}E_p u^i(f),$$

where u^i is a nonconstant affine vNM utility function over $\mathcal{X}$, $\mathcal{P}^i$ is a compact convex subset of Δ, and function $\alpha^i : B_0(\Sigma, u^i(\mathcal{X})) \to [0,1]$ reflects the extent of pessimism in the evaluation of acts (more precisely, utility profiles $u^i(f)$, $f \in \mathcal{F}$).[15] Moreover, u^i is unique up to positive affine transforms, $\mathcal{P}^i$ is unique, and α^i is constant on $\{k_1 h + k_2 : k_1 > 0, k_2 \in \mathbb{R}\}$ for any $h \in B_0(\Sigma, u^i(\mathcal{X}))$ and unique on utility profiles of non-crisp acts.[16] Without loss of generality, we set α^i to be $\frac{1}{2}$ for utility profiles of crisp acts. Notice that the model generalizes the Hurwicz α-pessimism criterion by allowing α^i to vary with acts.

GMM argues that the model disentangles ambiguity attitude, which is captured by α^i, from revealed ambiguity, which is captured by $\mathcal{P}^i$, and, ambiguity attitudes can only be compared among DMs who have the same revealed ambiguity. Under common revealed ambiguity, a more ambiguity averse DM is expected to put a higher weight on bad scenarios. The following characterization is thus quite natural.

Proposition 12. *Given generalized α-maxmin models $\{(u^i, \mathcal{P}^i, \alpha^i)\}_{i=1,2}$ with common revealed ambiguity $\mathcal{P}^1 = \mathcal{P}^2$, the following two statements are equivalent.*

1. *DM 1 is more ambiguity averse than DM 2.*
2. *$\alpha^1 \geq \alpha^2$.*

Intuitively the common revealed ambiguity condition might be satisfied in carefully designed lab environment or generally in situations when the DMs are likely to have the same information regarding uncertainty. Formally this condition can also be tested without the restriction of common risk attitude. To be specific, for any generalized α-maxmin models $\{(u^i, \mathcal{P}^i, \alpha^i)\}_{i=1,2}$, the corresponding unambiguous preferences $\{\succcurlyeq^{*i}\}_{i=1,2}$ (see GMM for definition) are Bewley preferences. Then in light of Corollary 4, for any four fixed prizes $a \succ^1 b$ and $c \succ^2 d$, we have $\mathcal{P}^1 = \mathcal{P}^2$ if and only if $f \succcurlyeq^{*1} \alpha a + (1-\alpha)b \Longleftrightarrow g \succcurlyeq^{*2} \alpha c + (1-\alpha)d$ for

[15] We define α^i over $B_0(\Sigma, u^i(\mathcal{X}))$ to streamline exposition, though originally GMM defines $\hat{\alpha}^i$ over $\mathcal{F}$. To construct α^i from $\hat{\alpha}^i$, for any $h \in B_0(\Sigma, u^i(\mathcal{X}))$ we define $\alpha^i(h) \equiv \hat{\alpha}^i(f)$ for any act f such that $u^i(f)$ is equal to some affine transform of h. By GMM's Lemma 8 and the property that $\hat{\alpha}^i$ is constant for acts with similar ambiguity (see GMM's Theorem 11), function α^i is well defined, and importantly, it is independent of the normalization of u^i.

[16] See GMM for the behavioral definition of crisp acts. In effect, an act f is crisp if and only if $min_{p\in\mathcal{P}^i}E_p u^i(f) = max_{p\in\mathcal{P}^i}E_p u^i(f)$. So, ambiguity does not affect crisp acts.

all aligned $f \in \mathcal{F}_{a,b}$ and $g \in \mathcal{F}_{c,d}$. This behavioral condition also generalizes GMM's in the spirit of our reformulation of comparative ambiguity attitudes.[17]

6.3 Homothetic Preferences

Chateauneuf and Faro (2009) axiomatically characterize homothetic preferences, according to which acts are evaluated as follows

$$f \succcurlyeq^i g \iff min_{p\in\Delta} \frac{1}{\varphi^i(p)} E_p u^i(f) \geq min_{p\in\Delta} \frac{1}{\varphi^i(p)} E_p u^i(f),$$

where u^i, unique up to a positive factor, is a nonconstant nonnegative affine vNM utility function over $\mathcal{X}$ with $u^i(\underline{m}) = 0$ for some least favorite prize $\underline{m}_i \in \mathcal{M}$, and $\varphi^i(p) = inf_{f\in\mathcal{F}}\big(\frac{E_p u^i(f)}{u^i(x_f)}\big) \in [0,1]$ is interpreted as the DM's confidence in prior $p \in \Delta$.[18] The SEU model can be seen as a special case, where the DM has confidence 1 in a single prior and confidence 0 in all other priors. A higher confidence function roughly corresponds to the DM being willing to consider more priors when evaluating acts, suggesting a more ambiguity averse attitude.

Because the least favorite prize plays a special role for any homothetic preference, we need to include it in the aligned acts used for comparing ambiguity attitudes. For this reason, we slightly adapt our formulation of ambiguity attitudes to this case.

Definition. Fix *any* two prizes $\{m_1, m_2\}$ such that $m_1 \succ^1 \underline{m}_1$ and $m_2 \succ^2 \underline{m}_2$. DM 1 is more ambiguity^{hp} averse than DM 2 if we have $\forall \alpha \in [0,1]$,

$$f \succcurlyeq^1 (m_1, \alpha; \underline{m}_1, 1-\alpha) \implies g \succcurlyeq^2 (m_2, \alpha; \underline{m}_2, 1-\alpha),$$

for all aligned $f \in \mathcal{F}_{m_1,\underline{m}_1}$ and $g \in \mathcal{F}_{m_2,\underline{m}_2}$.

In this formulation, the lesser prizes used in the aligned acts are fixed at the least favorite ones, while the better prizes are flexible. Therefore homothetic preferences are partially independent of the prizes used for comparison. A natural characterization follows.

[17] If one insists on comparing ambiguity attitudes when $\mathcal{P}^1 \neq \mathcal{P}^2$ (see Hartmann 2018), we suggest the following method that is still in the spirit of the likelihood approach. For each act $f \in \mathcal{F}_{m_1,m_2}$, there are a maximum $\underline{\alpha}^i_f$ and a minimum $\overline{\alpha}^i_f$ such that $\overline{\alpha}^i_f m_1 + (1-\overline{\alpha}^i_f s)m_2 \succcurlyeq^{*i} f \succcurlyeq^{*i} \underline{\alpha}^i_f m_1 + (1-\underline{\alpha}^i_f)m_2$ and an α^i_f such that $f \sim^i \alpha^i_f m_1 + (1-\alpha^i_f)m_2$. Then, it's clear that $\alpha^1 \geq \alpha^2$ is equivalent to $(\overline{\alpha}^1_f - \alpha^1_f)/(\overline{\alpha}^1_f - \underline{\alpha}^1_f) \geq (\overline{\alpha}^2_g - \alpha^2_g)/(\overline{\alpha}^2_g - \underline{\alpha}^2_g)$ for all aligned $f \in \mathcal{F}_{a,b}$ and $g \in \mathcal{F}_{c,d}$. However, it might be harder to carry out this comparison due to the indirect observability of $\succcurlyeq^{*i}$.

[18] In their Axiom 1, Chateauneuf and Faro (2009) assumes the existence of a least favorite consequence $x_* \in \mathcal{X}$, which should be a prize $\underline{m} \in \mathcal{M}$ in the Anscombe and Aumann setting.

Proposition 13. *Given homothetic preferences* $\{(u^i, \varphi^i)\}_{i=1,2}$*, the following two statements are equivalent.*

1. *DM 1 is more ambiguity*^{hp} *averse than DM 2.*
2. $\varphi^1 \geq \varphi^2$.

6.4 Uncertainty Averse Preferences

We now turn to the general class of uncertainty averse preferences à la Cerreia-Vioglio, Maccheroni, Marinacci, and Montrucchio (2011b, hereafter CMMM), which nest many popular models of decisions under uncertainty. This class of preferences is axiomatically characterized by CMMM , and it admits the following representation

$$f \succsim^i g \Leftrightarrow inf_{p\in\Delta} G^i\left(\int_{\mathcal{S}} u^i(f)dp, p\right) \geq inf_{p\in\Delta} G^i\left(\int_{\mathcal{S}} u^i(g)dp, p\right). \tag{6.1}$$

The function u^i is a nonconstant affine vNM utility function on $\mathcal{X}$, Δ is the set of (finitely additive) probability measures on $\mathcal{S}$, and the function $G^i : u^i(\mathcal{X}) \times \Delta \to (-\infty, +\infty]$ is quasiconvex on $u^i(\mathcal{X}) \times \Delta$, is increasing in its first argument, and satisfies $inf_{p\in\Delta} G^i(t, p) = t$ for all $t \in u(\mathcal{X})$. For lotteries, this model agrees with vNM expected utility. Moreover, though G^i is not unique in general, throughout this paper we follow CMMM to take the unique minimal G^i, which satisfies

$$G^i(t, p) = sup\left\{u^i(x_f) : E_p u^i(f) \leq t\right\}, \forall (t, p) \in u^i(\mathcal{X}) \times \Delta, \tag{6.2}$$

where for any $f \in \mathcal{F}$, $x_f \in \mathcal{X}$ is any lottery, whose existence is guaranteed by the axioms, such that $x_f \sim^i f$. The pair (u^i, G^i) is jointly unique in the following sense. If another pair $(\hat{u}^i, \hat{G}^i)$ also represent $\succsim$, it must be $\hat{u}^i = k_1 u^i + k_2$ for some $k_1 > 0$ and some $k_2 \in \mathbb{R}$, and $\hat{G}^i(t, p) = k_1 G^i(\frac{t}{k_1} - \frac{k_2}{k_1}, p) + k_2$ for all $(t, p) \in \hat{u}^i(\mathcal{X}) \times \Delta$.

Suppose DM $i \in \{1, 2\}$ has uncertainty averse preference $\succsim^i$ whose representation consists of components (u^i, G^i).

Theorem 14. *Given uncertainty averse preferences* $\{(u^i, G^i)\}_{i=1,2}$ *with the normalization* $u^1(\overline{m}_1) = u^2(\overline{m}_2)$ *and* $u^1(\underline{m}_1) = u^2(\underline{m}_2)$*, and under Assumption B, the following two statements are equivalent.*

1. *DM 1 is more ambiguity averse** *than DM 2.*
2. $G^1 \leq G^2$.

Let us discuss the rationale behind how $G^1 \leq G^2$ captures comparative ambiguity attitudes. For this purpose, it useful to draw an analogy between the function $G^i(t, p) =$

$sup\{u^i(x_f) : E_p u^i(f) \leq t\}$ and the indirect utility function of the Marshallian demand problem, for which t corresponds to a total expenditure, p corresponds to a price vector, $u^i(f)$ corresponds to a consumption bundle, and $u(x_f)$ corresponds to the utility level brought about by consumption bundle $u^i(f)$.[19] That is, a DM with budget t tries to allocate consumption goods across states, where each state s is associated with an Arrovian price $p(s)$. Since there always exists a lottery x such that $u^i(x) = t$, the constant, ambiguity-free bundle $u^i(x)$ is in the budget and it stands for the safest choice. Therefore, the size $G^i(t,p) - t \geq 0$ reflects the extent to which the DM is willing to "take a chance", for example, by choosing a bundle $u^i(f)$ that has higher quantity of goods on "cheaper" states. With this interpretation, $G^1 \leq G^2$ indicates that DM 1 is always more attracted to safer choices, hence a more ambiguity averse attitude. Viewed from this perspective, if one were to ask, when is it the case that $G^1 \leq G^2$, our theorem provides a characterization of this inequality that holds on a significantly richer domain (of pairs of preference relations) than that of CMMM.

Now we turn to two well known special cases of uncertainty averse preferences: variational preferences and multiplier preferences.

6.4.1 Variational Preferences

Maccheroni, Marinacci, and Rustichini (2006) axiomatize variational preferences, which evaluate acts as

$$f \succsim^i g \iff min_{p\in\Delta}\{E_p u^i(f) + c^i(p)\} \geq min_{p\in\Delta}\{E_p u^i(g) + c^i(p)\},$$

where u^i is a nonconstant affine vNM utility function over $\mathcal{X}$, and $c^i : p \in \Delta \mapsto sup_{f\in\mathcal{F}}\{u^i(x_f) - E_p u^i(f)\} \in [0, +\infty]$, where for every act f, x_f is a lottery such that $f \sim^i x_f$. If another pair $(\hat{u}^i, \hat{c}^i)$ represents the same preference, it must be the case that $\hat{u}^i = k_1 u^i + k_2$ for some $k_1 > 0$ and $k_2 \in \mathbb{R}$, and $\hat{c}^i = k_1 c^i$. Function $c^i(\cdot)$ reflects the subjective cost of using each prior, as in Akaike's (1974) information criterion. A lower cost function roughly corresponds to the DM willing to consider more priors when evaluating acts, suggesting a more ambiguity averse attitude.

Corollary 15. *Given variational preferences $\{(u^i, c^i)\}_{i=1,2}$ with the normalization $u^1(\overline{m}_1) = u^2(\overline{m}_2)$ and $u^1(\underline{m}_1) = u^2(\underline{m}_2)$, and under Assumption B, the following two statements are equivalent.*

[19] Indeed, the relation between equation (6.1) and (6.2) is formally similar to that between the utility function and the indirect utility function in the Marshallian demand problem. See Cerreia-Vioglio, Maccheroni, Marinacci, and Montrucchio (2011a) for a study of dualities of this kind.

1. *DM 1 is more ambiguity averse* than DM 2.*
2. $c^1 \leq c^2$.

6.4.2 Multiplier Preferences

Strzalecki (2011) axiomatizes a popular special case of variational preferences named multiplier preferences, where the cost function takes the form of relative Shanon entropy (see also, Hansen and Sargent 2008). Specifically, acts are evaluated as

$$f \succsim^i g \iff min_{p\in\Delta}\{E_p u^i(f) + \theta^i R(p||q^i)\} \geq min_{p\in\Delta}\{E_p u^i(g) + \theta^i R(p||q^i)\},$$

where u^i is a nonconstant affine vNM utility function over $\mathcal{X}$, $\theta^i \in (0,\infty]$ is a multiplier, $q^i \in \Delta$ is a unique σ-additive probability, and $R(p||q^i)$ is the relative entropy between p and q^i.[20] A lower θ^i indicates that the DM is willing to consider more priors when evaluating acts, hence a more ambiguity averse attitude.

Corollary 16. *Given multiplier preferences $\{(u^i,\theta^i,q^i)\}_{i=1,2}$ with the normalization $u^1(\overline{m}_1) = u^2(\overline{m}_2)$ and $u^1(\underline{m}_1) = u^2(\underline{m}_2)$, and under Assumption B, the following two statements are equivalent.*

1. *DM 1 is more ambiguity averse* than DM 2.*
2. *$\theta^1 \leq \theta^2$ and $q^1 = q^2$.*

Notice that the applicability of the comparison implies that the DMs share the same base probability $q^1 = q^2$.

7 Discussion

7.1 Linear Utility Functions

As we point out in the explanation of condition (2.5), the linear structure of vNM utility enables using objective probabilities to cardinally measure the desirability of ambiguous acts. Linear perception of probabilities therefore underlies the simplicity and strength of our results. However, since Allais (1953), there is ample evidence that people's perception of probabilities may be nonlinear (see, for example, Gonzalez and Wu 1999). In spite of this challenge, our formulation of ambiguity attitudes might be valid without modification if there are only two states of the world (see Dimmock, Kouwenberg, and Wakker 2016 and

[20] Recall that the relative entropy $R(p||q^i)$ between p and q^i is $\int_{\mathcal{S}}(log\frac{dp}{dq^i})dp$ if p is σ-additive and absolutely continuous to q^i and $+\infty$ otherwise.

Wang 2019). And, in general, the main idea might still work if we state condition (2.5) using each DM's "subjective probability scale", which can be derived from observable behaviors. For example, under the assumption that the DMs abide by the rank-dependent utility model over lotteries (Quiggin 1982), Wang (2019) utilizes a notion of subjective probabilistic midpoints to formulate comparative ambiguity attitudes under maxmin preferences.[21]

7.2 Assumption B

The discussion here is conducted in the context of prize-dependent models. Prize-independent models do not need Assumption B and they are free from the subtleties discussed below.

Foremost, Assumption B is a technical condition that allows us to study all the aspects of a DM's ambiguity attitude using two prizes. Specifically, with $u^i(\mathcal{M}) = [u^i(\overline{m}_i), u^i(\underline{m}_i)]$, for any $f \in \mathcal{F}$ we can find a $f' \in \mathcal{F}_{\overline{m}_i,\underline{m}_i}$ such that $u(f(s)) = u(f'(s))$ for all s. Therefore, $\mathcal{F}_{\overline{m}_1,\underline{m}_1}$ and $\mathcal{F}_{\overline{m}_2,\underline{m}_2}$ are rich enough to cover the whole range of issues that the DMs may consider.

This naturally gives condition (2.5) a flavor of *pointwise* comparison. Namely, for any utility profile ψ, which is a simple function from S to the normalized range of utility $[u^1(\underline{m}_1), u^1(\overline{m}_1)] = [u^2(\underline{m}_2), u^2(\overline{m}_2)]$, we can find aligned acts $f \in \mathcal{F}_{\overline{m}_1,\underline{m}_1}$ and $g \in \mathcal{F}_{\overline{m}_2,\underline{m}_2}$ such that $u^1(f) = u^2(g) = \psi$. Condition (2.5) can be understood as saying that DM 1 is more ambiguity averse than DM 2 at every utility profile.[22] The prize-dependence nature indicates that, when condition (2.5) holds for aligned acts on $\mathcal{F}_{\overline{m}_1,\underline{m}_1}$ and $\mathcal{F}_{\overline{m}_2,\underline{m}_2}$, it might not hold for aligned acts on $\mathcal{F}_{a,b}$ and $\mathcal{F}_{c,d}$ for some other prizes $a \succ^1 b$ and $c \succ^2 d$. This might be the case, for example, when $[u^1(b), u^1(a)] = [100u^2(c), 100u^2(d)] \subset \mathbb{R}_{++}$, because for every pair of aligned acts $f \in \mathcal{F}_{a,b}$ and $g \in \mathcal{F}_{c,d}$ we have $u^1(f(s)) = 100u^2(g(s)) > 0$ for all s, that is, f looks much better to DM 1 than does g to DM 2. So, we might have DM 1 being more ambiguity averse than DM 2 on $\mathcal{F}_{\overline{m}_1,\underline{m}_1}$ and $\mathcal{F}_{\overline{m}_2,\underline{m}_2}$ but less ambiguity averse on $\mathcal{F}_{a,b}$ and $\mathcal{F}_{c,d}$, especially when DMs have decreasing absolute ambiguity attitude. However, when utility ranges $[u^1(b), u^1(a)]$ and $[u^2(d), u^2(c)]$ coincide, a condition that can be verified in principle (e.g., by deriving utility mid-points from behavior, see Abdellaoui 2000), such dependence is circumvented and therefore condition (2.5) also holds for aligned acts on $\mathcal{F}_{a,b}$ and $\mathcal{F}_{c,d}$.

[21] For any two probability numbers $\alpha, \beta \in [0,1]$, Wang (2019) finds, via observing finitely many preference relations of a DM, the probability number γ that is subjectively perceived by the DM to be the midpoint of α and β in terms of probability weighting.

[22] It is indeed well defined to talk about comparative ambiguity attitude at a utility profile. Specifically, fix any utility profile ψ, and suppose there are aligned acts $f' \in \mathcal{F}_{a,b}$ and $g' \in \mathcal{F}_{c,d}$ with $u^1(f') = u^2(g') = \psi$ for for some other prizes $a \succ^1 b$ and $c \succ^2 d$. This in general indicates $[u^1(b), u^1(a)] = [u^2(d), u^2(c)]$. And, importantly, if DM 1 is more ambiguity averse than DM 2 under $\mathcal{F}_{\overline{m}_1,\underline{m}_1}$ and $\mathcal{F}_{\overline{m}_2,\underline{m}_2}$, we will also have $mp^1(f') \leq mp^2(g')$.

There might be two potential ways to relax Assumption B. Take $\mathcal{M}$ being the set $\mathbb{R}$ of monetary prizes for example. The first approach requires that, for any two ranges of prizes $[m_1, M_1], [m_2, M_2] \subset \mathbb{R}$, DM 1 is more ambiguity averse than DM 2 for aligned acts on $\mathcal{F}_{M_1,m_1}$ and $\mathcal{F}_{M_2,m_2}$. This condition has the flavor of DM 1 being *uniformly* more ambiguous averse than DM 2. Such a requirement might turn out to be too strong. Take, for example, smooth ambiguity models $\{(\phi^i, u^i, \mu^i)\}_{i=1,2}$ and ϕ^1 is concave and continuously differentiable. When $[u^1(m_1), u^1(M_1)]$ is very narrow, DM 1 is almost ambiguity neutral when evaluating acts in $\mathcal{F}_{M_1,m_1}$ (see Lang 2017). Therefore, requiring DM 1 to be uniformly more ambiguity averse is likely to force DM 2 to be ambiguity neutral or ambiguity seeking.

The second approach requires that for any *common* range of prizes $[m, M] \subset \mathbb{R}$, DM 1 is more ambiguity averse than DM 2 for all aligned acts $f \in \mathcal{F}_{M,m}$ and $g = f \in \mathcal{F}_{M,m}$. The corresponding formal characterization will be more complicated, which roughly amounts to requiring our characterizations to hold on utility ranges $[u^1(m), u^1(M)]$ and $[u^2(m), u^2(M)]$ for all $m < M$. This approach is taken by Cerreia-Vioglio, Maccheroni, and Marinacci (2019) and also by this paper to study absolute ambiguity attitudes, where the two DMs are actually the same DM at two base levels of wealth. However, for the genuine inter-personal comparisons discussed in section 6, we choose to adopt Assumption B for a simpler exposition and closer ties to previous results in the literature.

7.3 Structure of the Set of Consequences

Our formulation of comparative ambiguity attitudes relies on the apparatus of objective probabilities (or mixing coefficients in general), which is readily available in many axiomatic decision theoretic models. For empirical purposes, a researcher can always introduce objective probabilities (or events with known probabilities) to facilitate the comparison. If we have to do without objective probabilities, some strategies might still work in relaxing the common-preference restriction. Here we discuss two of them, which correspond to two types of structures on the set $\mathcal{X}$ of consequences.

Epstein (1999) and Epstein and Zhang (2001) propose, in a Savagean setting where $\mathcal{X}$ has no structure, a definition of ambiguity aversion using probability sophistication (Machina and Schmeidler 1992) as the benchmark of ambiguity neutrality.[23] Trying to achieve a separation between risk attitude and ambiguity attitude, they assume the existence of a collection Σ_{ua} of unambiguous events, on which risk attitude resides. A DM with preference $\succcurlyeq$ is ambiguity

[23] Roughly, a probability sophisticated DM has a unique prior relative to which acts are reduced to lotteries, which are then evaluated according to the DM's preference over lotteries.

averse if there is a probability sophisticated DM with preference $\succcurlyeq^{ps}$ such that

$$h \succcurlyeq^{ps} (\succ^{ps}) f \implies h \succcurlyeq (\succ) f \tag{7.1}$$

where f is any act and h is any unambiguous act (that is, act measurable to Σ_{ua}). Condition (7.1) implies that the two DMs share preferences over unambiguous acts and therefore the same risk attitude[24], which is then used as a reference system to indicate that the ambiguous averse DM always has a weakly lower evaluation of acts. Following this paper's idea, we are naturally led to using unambiguous events themselves as the reference system to measure the desirability of acts. Thus we have the following variation of condition (7.1): for some $a \succ^{ps} b$ and $c \succ d$,

$$(a, E; b, E^c) \succcurlyeq^{ps} (\succ^{ps}) f \implies (c, E; d, E^c) \succcurlyeq (\succ) g \tag{7.2}$$

for any unambiguous event E and any pair of acts f and g such that for all $s \in \mathcal{S}$, $f(s) \in \{a, b\}$, $g(s) \in \{c, d\}$, and $f(s) = a$ if and only if $g(s) = c$. This variation allows the preference over $\mathcal{X}$ to vary across the DMs. However, it might be generally not strong enough to deliver formal characterizations, because, first, the risk attitude allowed by Epstein (1999) and Epstein and Zhang (2001) is more general than that of expected utility, and, second, the lack of structure on $\mathcal{X}$ renders the set of pairs of comparable acts f and g in condition (7.2) quite meagre.[25] However, it might serve as an inspiration for statistical tests or future work on more structured settings or preferences.

Another popular structure on $\mathcal{X}$ is topological connectedness. This is employed by, for example, Wakker (1989), in the derivation of Choquet expected utility, and, Alon and Schmeidler (2014), in the derivation of maxmin expected utility. Authors of these works usually start by studying tradeoffs among consequences, which are manifested in observable preference relations. Such effort will eventually establish, for every DM $i \in \{1, 2\}$, every two consequences $x, y \in \mathcal{X}$, and every dyadic number $r \in [0, 1]$, the existence of a consequence z that is the *subjective $r-$mixture* between x and y, denoted by $rx \oplus^i (1-r)y$. In effect, when $z \sim^i rx \oplus^i (1-r)y$, it satisfies $u^i(z) = ru^i(x) + (1-r)u^i(y)$, where u^i is DM i's utility function over $\mathcal{X}$. Naturally, we can restate the formulations of comparative ambiguity attitudes in this paper by replacing lotteries with subjective mixtures among consequences. We conjecture that the characterizations obtained here would have natural counterparts in these models as well.

[24]See Epstein (1999) and Epstein and Zhang (2001) for discussions. Notice that common risk attitude include common utility function and all other components (eg. probability weighting) associated with risk attitude.

[25]A formal characterization of comparative ambiguity attitudes may be obtainable for Choquet expected utility under the setting of Epstein and Zhang (2001), because the collection Σ_{ua} therein is rich in the sense that the (unambiguous) probabilities on Σ_{ua} is convex ranged.

8 Appendix

8.1 Proofs

Proof of Theorem 3

Proof. Since $\mathcal{P}^i$ is unique and independent of u^i, we may take the normalization $u^1(a) = u^2(c) = 1$ and $u^1(b) = u^2(d) = 0$.

(1) implies (2): If there is a $p \in \mathcal{P}^2$ such that $p \notin \mathcal{P}^1$, by a hyperplane separation argument there is a simple function $l : \mathcal{S} \to \mathbb{R}$ such that $min_{p' \in \mathcal{P}^1} E_{p'} l > E_p l \geq min_{p' \in \mathcal{P}^2} E_{p'} l$. Since these inequalities are unaffected by positive affine transforms of l, we can without loss of generality assume $0 \leq l(s) \leq 1, \forall s \in \mathcal{S}$. Then there are aligned acts $f \in \mathcal{F}_{(a,b)}$ and $g \in \mathcal{F}_{(c,d)}$ such that $\{u^1(f(s))\}_{s \in \mathcal{S}} = \{u^2(g(s))\}_{s \in \mathcal{S}} = l$. Then we have $min_{p' \in \mathcal{P}^1} E_{p'} u^1(f) > min_{p' \in \mathcal{P}^2} E_{p'} u^2(g)$ and therefore (1) has to be violated .

(2) implies (1): If acts $f \in \mathcal{F}_{(a,b)}$ and $g \in \mathcal{F}_{(c,d)}$ are aligned, we have $\{u^1(f(s))\}_{s \in \mathcal{S}} = \{u^2(g(s))\}_{s \in \mathcal{S}}$. Then, $\mathcal{P}^1 \supset \mathcal{P}^2$ indicates $min_{p' \in \mathcal{P}^1} E_{p'} u^1(f) \leq min_{p' \in \mathcal{P}^2} E_{p'} u^2(g)$, hence (1). $\square$ $\square$

Proof of Corollary 4

Proof. Take the normalization $u^1(a) = u^2(c) = 1$ and $u^1(b) = u^2(d) = 0$. For any $\psi \in B(\Sigma, [0,1])$, (1) is equivalent to

$$min_{p' \in \mathcal{P}^1} E_{p'} \psi \leq min_{p' \in \mathcal{P}^2} E_{p'} \psi, \forall \psi \in B(\Sigma, [0,1]).$$

From here the proof of Theorem 3 can be reused. $\square$

Proof of Theorem 6

Proof. Without loss of generality we assume $u^1(\overline{m}_1) = u^2(\overline{m}_2) = 1$ and $u^1(\underline{m}_1) = u^2(\underline{m}_2) = 0$. Define $\psi \equiv \phi^1 \circ (\phi^2)^{-1} : \phi^2([0,1]) \to \mathbb{R}$, which is continuous and strictly increasing. (2) is true if and only if ψ is concave.

(1) implies (2): This result will be established in two steps.

In step 1, we turn to the part $\mu^1 = \mu^2$. To prove it, we will show that μ^1 and μ^2 share the same "mean" first order belief, that is $\int_\Delta p(E) d\mu^1(p) = \int_\Delta p(E) d\mu^2(p)$ for all $E \in \Sigma$, which in combination with Assumption CS imply $\mu^1 = \mu^2$.[26] Suppose $\int_\Delta p d\mu^1 \neq \int_\Delta p d\mu^2$, then there exists $h \in B_0(\Sigma, \mathbb{R})$ such that $\int_\Delta E_p h d\mu^1(p) > \int_\Delta E_p h d\mu^2(p)$. By translation and

[26]Notice that, for every $E \in \Sigma$, the function that maps every $p \in \Delta$ to $p(E)$ is weak* continuous and henceforth measurable.

scaling we can without loss of generality take $h \in B_0(\Sigma,[0,+\infty))$ such that $\int_\Delta E_p h d\mu^1(p) = \int_\Delta E_p h d\mu^2(p) + 1$. Suppose ϕ^1 and ϕ^2 is continuously differentiable on some interval $[d, d+\epsilon] \subset [0,1]$. For $\delta > 0$ small enough we have $d + \delta h \in B_0(\Sigma,[d, d+\epsilon])$, which is the utility profile of some aligned acts $f \in \mathcal{F}_{\overline{m}_1,\underline{m}_1}$ and $g \in \mathcal{F}_{\overline{m}_2,\underline{m}_2}$. Then, statement (1) implies that for all small enough $\delta > 0$,

$$(\phi^1)^{-1}\Big[\int \phi^1\big(E_p(d+\delta h)\big)d\mu^1(p)\Big] \leq (\phi^2)^{-1}\Big[\int \phi^2\big(E_p(d+\delta h)\big)d\mu^2(p)\Big] \tag{8.1}$$

However, inequality (8.1) must be violated for some very small $\delta > 0$ because ϕ^1 and ϕ^2 are locally linear on very small intervals in $[d, d+\epsilon]$, indicating that $\int E_p(d+\delta h)d\mu^1(p) > \int E_p(d+\delta h)d\mu^2(p)$ will be preserved even after transformations. Rigorously, by Taylor expansion around $\int E_p(d+\delta h)d\mu^i(p)$ to the first degree, we are led to

$$\int \phi^i\big(E_p(d+\delta h)\big)d\mu^i(p) = \phi^i(d + \delta \int E_p h d\mu^i) + \int \big((\phi^i)'(d) + o(1)\big)\delta\big(h - \int E_p h d\mu^i\big)d\mu^i$$

and therefore the following two inequalities

$$\begin{aligned}
(\phi^1)^{-1}\Big[\int \phi^1\big(E_p(d+\delta h)\big)d\mu^1(p)\Big] &\geq d + \delta + \delta \int E_p h d\mu^2(p) - \left|\frac{o(\delta)}{\phi_1'(d) - o(1)}\right|; \\
(\phi^2)^{-1}\Big[\int \phi^2\big(E_p(d+\delta h)\big)d\mu^2(p)\Big] &\leq d + \delta \int E_p h d\mu^2(p) + \left|\frac{o(\delta)}{\phi_2'(d) - o(1)}\right|,
\end{aligned}$$

where, as $\delta \downarrow 0$, $o(1) \to 0$ and $o(\delta)/\delta \to 0$. Hence we have

$$(\phi^1)^{-1}\Big[\int \phi^1\big(E_p(d+\delta h)\big)d\mu^1(p)\Big] - (\phi^2)^{-1}\Big[\int \phi^2\big(E_p(d+\delta h)\big)d\mu^2(p)\Big] \geq \delta - o(\delta).$$

As $\delta \downarrow 0$, this difference is eventually going to be strictly positive, which contradicts inequality (8.1). So we must have $\int_\Delta p d\mu^1 = \int_\Delta p d\mu^2$. In addition, since Assumption CS implies that each $q \in conv(\{q_n\}_{n=1}^N)$ corresponds to an unique combination of weights on $\{q_n\}_{n=1}^N$, we conclude that $\mu^1 = \mu^2$.

In step 2, we turn to the part $\phi^1 = \psi \circ \phi^2$ for some concave function $\psi : \phi^2(u^2(\mathcal{X})) \to \mathbb{R}$. If $\mu^1 = \mu^2$ has a support that consists of a single $q \in \{q_n\}_{n=1}^N$ we take ϕ^1, ϕ^2, and ψ to be the identity function. Otherwise, $\mu^1 = \mu^2$ has a nontrivial support. Then fix any $q \in \{q_n\}_{n=1}^N$ such that $\mu^1(q) = \mu^2(q) = \gamma \in (0,1)$. Since q is disjoint from the subspace spanned by $\{q_n\}_{n=1}^N \backslash \{q\}$, there is a $l \in B_0(\Sigma,\mathbb{R})$ such that $E_q l > 0$ and $E_{q'} l = 0$ for all $q' \in \{q_n\}_{n=1}^N \backslash \{q\}$. By scaling we can without loss of generality take $l \in B_0(\Sigma,[\underline{l},\overline{l}])$ such that $E_q l = 1$ and $E_{q'} l = 0$ for all $q' \in \{q_n\}_{n=1}^N \backslash \{q\}$, where $\underline{l} = min_s l(s) \leq 0$ and $\overline{l} = max_s l(s) \geq 1$.

For all $a \in (0,1)$, define $\Lambda(a) = min(\frac{a}{|\underline{l}|}, \frac{1-a}{\overline{l}}) > 0$, for which $\frac{a}{|\underline{l}|}$ is taken to be $+\infty$ if

$\underline{l} = 0$. Observe that for all $a \in (0,1)$ and all $\lambda \in (0, \Lambda(a))$, we have $a + \lambda \cdot l \in B_0(\Sigma, [0,1])$, which therefore is the utility profile of some aligned acts $f \in \mathcal{F}_{\overline{m}_1, \underline{m}_1}$ and $g \in \mathcal{F}_{\overline{m}_2, \underline{m}_2}$. Since $E_q(a + \lambda \cdot l) = a + \lambda$ and $E_{q'}(a + \lambda \cdot l) = a$ for all $q' \in \{q_n\}_{n=1}^N \backslash \{q\}$, then for this pair of aligned f and g statement (1) implies

$$(\phi^1)^{-1}\big[\gamma\phi^1(\alpha + \lambda) + (1-\gamma)\phi^1(\alpha)\big] \leq (\phi^2)^{-1}\big[\gamma\phi^2(\alpha + \lambda) + \gamma\phi^2(\alpha)\big].$$

Since $a + \Lambda(a)$ weakly increases in $a \in (0,1)$, for any $a' \in (a, a + \Lambda(a))$ and any $\lambda' > 0$ such that $a' + \lambda' < a + \Lambda(a)$ we must have $\lambda' < \Lambda(a')$ and therefore $a' + \lambda' \cdot l \in B_0(\Sigma, [0,1])$. Since $a' + \lambda' \cdot l$ is also the utility profile of some aligned acts $f \in \mathcal{F}_{\overline{m}_1, \underline{m}_1}$ and $g \in \mathcal{F}_{\overline{m}_2, \underline{m}_2}$, and by the above inequality, we have for all $a', a' + \lambda' \in [a, a + \Lambda(a))$ and $\lambda' > 0$,

$$\gamma\psi \circ \phi^2(\alpha' + \lambda') + (1-\gamma)\psi \circ \phi^2(\alpha') \leq \psi\big[\gamma\phi^2(\alpha' + \lambda') + (1-\gamma)\phi^2(\alpha')\big], \tag{8.2}$$

which can be reformulated as for all $b, c \in \phi^2([a, a + \Lambda(a)))$ and $b < c$,

$$\gamma\psi(c) + (1-\gamma)\psi(b) \leq \psi\big[\gamma c + (1-\gamma)b\big]. \tag{8.3}$$

The inequality (8.3) implies that ψ is concave on $\phi^2((a, a+\Lambda(a)))$. For a simple argument, suppose $\psi(\alpha y + (1-\alpha)x) < \alpha\psi(y) + (1-\alpha)\psi(x)$ for some $\alpha \in (0,1)$ and $x < y \in \phi^2((a, a + \Lambda(a)))$. Consider the line segment l_{xy} connecting $(x, \psi(x))$ to $(y, \psi(y))$, and, because ψ is continuous, take $b < \alpha y + (1-\alpha)x$ to be the largest value such that $(b, \psi(b))$ is on l_{xy} and $c > \alpha y + (1-\alpha)x$ to be the smallest value such that $(c, \psi(c))$ is on l_{xy} . Then, ψ restricted on (b, c) must be strictly below l_{xy} restricted on (b, c), which contradicts inequality (8.3). The claim is thus valid.

Since ψ is concave on $\phi^2((a, a+\Lambda(a)))$ for all $a \in (0,1)$, ψ must be concave on $\phi^2([0,1])$. To see why, consider any $0 < a < a' < 1$ such that $(a, a+\Lambda(a))$ and $(a', a+\Lambda(a'))$ have nonempty intersection and $a+\Lambda(a) < a'+\Lambda(a')$. For any $b \in (a, a+\Lambda(a))$ and $c \in (a+\Lambda(a), a'+\Lambda(a'))$, notice that the line segment connecting $(\phi^2(b), \psi(\phi^2(b)))$ to $(\phi^2(c), \psi(\phi^2(c)))$ is below the line segments connecting $(\phi^2(b), \psi(\phi^2(b)))$ first to $(\phi^2(a + \Lambda(a)), \psi(\phi^2(a + \Lambda(a))))$ and then to $(\phi^2(c), \psi(\phi^2(c)))$, which are still below ψ. This indicates ψ is concave on $\phi^2((a, a' + \Lambda(a')))$. Applying this argument consecutively, we get ψ being concave on $\phi^2((0,1))$. Then, ψ is concave on $\phi^2([0,1])$ by continuity.

(2) implies (1): Any pair of aligned acts $f \in \mathcal{F}_{\overline{m}_1, \underline{m}_1}$ and $g \in \mathcal{F}_{\overline{m}_2, \underline{m}_2}$ induce the same utility profile $u^1(f)$ and $u^2(g)$, which along with $\mu^1 = \mu^2$ imply $\mu^1(\{p : E_p u^1(f) \leq a\}) = \mu^2(\{p : E_p u^2(g) \leq a\})$ for all $a \in \mathbb{R}$. Because ϕ^1 is more concave, we have $(\phi^1)^{-1}[E_{\mu^1}\phi^1(E_p u^1(f))] \leq (\phi^2)^{-1}[E_{\mu^2}\phi^2(E_p u^2(g))]$, which means the risk equivalent of f is weakly lower than that of g,

and, henceforth, statement (1) is true. $\square$

Proof of Proposition 11

Proof. We show (1)⇒(2)⇒(3)⇒(1).

(1)⇒(2): This obvious because (2) is a special case of (1).

(2)⇒(3): Notice that $\forall E$, $(a, E; b, E^c) \sim^1 (a, p^1_E; b, E^c)$ is equivalent to $v^1(E) = p^1_E$. DM 2's case is similar. So, $\forall E$, $p^1_E \leq p^2_E$ indicates $v^1(E) \leq v^2(E)$.

(3)⇒(1): Given $u^1(a) = u^2(c) = 1$ and $u^1(b) = u^2(d) = 0$, (3) means that for any pair of aligned acts f and g, DM 1's prospect is first order dominated by DM 2's prospect. This means DM 1 has a lower evaluation and therefore (1). $\square$

Proof of Proposition 12

Proof. Notice the following two facts. First, $\forall h \in B_0$, α^i is by construction (again, according to GMM's Lemma 8 and Theorem 11) constant on $A(h) \equiv \{k_1 h + k_2 : k_1 > 0, k_2 \in \mathbb{R}\}$. Second, $\forall h \in B_0$, we can find aligned acts $f \in \mathcal{F}_{(a,b)}$ and $g \in \mathcal{F}_{(c,d)}$ such that $u^1(f), u^2(g) \in A(h)$, and for any aligned acts $f \in \mathcal{F}_{(a,b)}$ and $g \in \mathcal{F}_{(c,d)}$, $\exists h \in B_0$ such that $u^1(f), u^2(g) \in A(h)$.

(1)⇒(2): For any $f \in \mathcal{F}_{(a,b)}$ and $g \in \mathcal{F}_{(c,d)}$ that are aligned, some simple algebra (1) lead to $\alpha^1[u^1(f)] \geq \alpha^2[u^2(g)]$. This combined with the two facts lead to $\alpha^1 \geq \alpha^2$.

(2)⇒(1): If $f \in \mathcal{F}_{(a,b)}$ and $g \in \mathcal{F}_{(c,d)}$ are aligned, $\alpha^1 \geq \alpha^2$ combined with the two facts imply that $\alpha^1[u^1(f)] \geq \alpha^2[u^2(g)]$, which further implies (1). $\square$

Proof of Proposition 13

Proof. Since φ^i is unique and thus independent from u^i, we can without loss of generality normalize $u^1(m_1) = u^2(m_2) = 1$. Recall that $u^1(\underline{m}_1) = u^2(\underline{m}_2) = 0$.

We claim $\varphi^1(p) = inf_{f \in \mathcal{F}}\left(\frac{E_p u^1(f)}{u^1(x^1_f)}\right) = inf_{f \in \mathcal{F}_{(m,\underline{m}_1)}}\left(\frac{E_p u^1(f)}{u^1(x^1_f)}\right)$, $\forall p \in \Delta$. For any $f \in \mathcal{F}$ we can find a $k > 0$ such that there is a $f' \in \mathcal{F}_{(m_1,\underline{m}_1)}$ and $u^1(f(s)) = k u^1(f'(s))$ for all $s \in \mathcal{S}$. As $min_{p \in \Delta} \frac{1}{\varphi^1(p)} E_p u^1(f) = k \cdot min_{p \in \Delta} \frac{1}{\varphi^1(p)} E_p u^1(f')$, we have $u^1(x^1_f) = k \cdot u^1(x^1_{f'})$. So $\frac{E_p u^1(f)}{u^1(x^1_f)} = \frac{E_p u^1(f')}{u^1(x^1_{f'})}$, which proves the claim.

Similarly we have $\varphi^2(p) = inf_{g \in \mathcal{F}_{(m_2,\underline{m}_2)}}\left(\frac{E_p u^2(f)}{u^2(x^2_g)}\right)$, $\forall p \in \Delta$.

(1) implies (2): For any aligned pair of acts $f \in \mathcal{F}_{(m_1,\underline{m}_1)}$ and $g \in \mathcal{F}_{(m_2,\underline{m}_2)}$, (1) is equivalent to $u^1(x^1_f) \leq u^2(x^2_g)$, where $f \sim^1 x^1_f \in \mathcal{X}$ and $g \sim^2 x^2_g \in \mathcal{X}$. This leads to $\frac{E_p u^1(f)}{u^1(x^1_f)} \geq \frac{E_p u^1(g)}{u^2(x^2_g)}$ for all p, which implies $\varphi^1 \geq \varphi^2$.

(2) implies (1): (2) implies $min_{p \in \Delta} \frac{1}{\varphi^1(p)} E_p u^1(f) \leq min_{p \in \Delta} \frac{1}{\varphi^2(p)} E_p u^2(f)$, for any aligned pair of acts $f \in \mathcal{F}_{(m_1,\underline{m}_1)}$ and $g \in \mathcal{F}_{(m_2,\underline{m}_2)}$. That means $u^1(x^1_f) \leq u^2(x^2_g)$, leading to (1). $\square$

Proof of Theorem 14

Proof. We assume the normalization $u^1(\overline{m}_1) = u^2(\overline{m}_2)$ and $u^1(\underline{m}_1) = u^2(\underline{m}_2)$ throughout the proof. Notice that for any $f \in \mathcal{F}$ and any $s \in \mathcal{S}$, we can find an $\alpha^i \in [0,1]$ such that $f(s) \sim^i \alpha^i \overline{m}_i + (1-\alpha^i)\underline{m}_i$ since $\overline{m}_i \succcurlyeq^i f(s) \succcurlyeq^i \underline{m}_i$. Then, by monotonicity, we can always find a $f^i \in \mathcal{F}_{(\overline{m}_i,\underline{m}_i)}$ such that $f \sim^i f^i$. So it is without loss of generality to restrict attention on $\mathcal{F}_{(\overline{m}_i,\underline{m}_i)}$.

(1) implies (2): (1) implies that if f and g are aligned it must be the case that $u^1(x^1_f) \leq u^2(x^2_g)$, where x^1_f and x^2_g are consequences such that $x^1_f \sim^1 f$ and $x^2_g \sim^2 g$. Now, recall that $G^i(t,p) = sup\left\{u^i(x_f) : E_p u^i(f) \leq t\right\}$. For any $g \in \mathcal{F}_{(\overline{m}_2,\underline{m}_2)}$ with $E_p u^2(g) \leq t$, we can find an aligned $f \in \mathcal{F}_{(\overline{m}_1,\underline{m}_1)}$, with $u^1(f(s)) = u^2(g(s))$ for all $s \in \mathcal{S}$ and thus $\int u^1(f)dp \leq t$. Since $u^1(x^1_f) \leq u^2(x^2_g)$, it must be true that $G^1 \leq G^2$.

(2) implies (1): Since f and g are aligned, we have $\int u^1(f)dp = \int u^2(g)dp$ for all $p \in \Delta$. Then, $G^1 \leq G^2$ gives us $inf_{p\in\Delta} G^1\left(\int u^1(f)dp, p\right) \leq inf_{p\in\Delta} G^2\left(\int u^2(g)dp, p\right)$. Combine this inequality with the fact that there are $\alpha^1_f, \alpha^2_g \in [0,1]$ such that $f \sim^1 \alpha^1_f \overline{x}_1 + (1-\alpha^1_f)\underline{x}_1$ and $g \sim^2 \alpha^2_g \overline{x}_2 + (1-\alpha^2_g)\underline{x}_2$, we can conclude that $\alpha^1_f \leq \alpha^2_g$, which implies (1). $\square$

Proof of Corollary 15

Proof. A variational preference (u^i, c^i) is an uncertainty averse preferences with $G^i(t,p) = t + c^i(p)$. The result is then a simple corollary of Theorem 14. $\square$

Proof of Corollary 16

Proof. (1) implies (2): By Corollary (15) and the fact that $c^i(p) = \theta^i R(p||q^i)$ for multiplier preference (u^i, θ^i, q^i), we have $\theta^1 R(p||q^1) \leq \theta^2 R(p||q^2)$ for all $p \in \Delta$. Then, for $p = q^2$ we have $\theta^1 R(q^2||q^1) \leq \theta^2 R(q^2||q^2) = 0$, which means $R(q^2||q^1) = 0$ and thus $q^1 = q^2$. Then, $\theta^1 \leq \theta^2$ is immediate.

(2) implies (1): This follows from Corollary (15). $\square$

Proof of Theorem 9

Proof. The equivalence between statement 1 and 2 follows directly from the definition of DAAA and monotonicity.

We show statement 2 and 3 are equivalent. Take any aligned $f \in \mathcal{F}_{v+a,v}$ and $g \in \mathcal{F}_{w+a,w}$, we have $u(f) = u(v) + \psi$ for some $\psi \in B_0(\Sigma, [0, u(v+a) - u(v)])$ and $u(g) = u(w) + \lambda\psi$ with $\lambda = [u(w+a) - u(w)]/[u(v+a) - u(v)]$. Then by the definition of risk equivalent, we have $I(u(v)+\psi) = u(v) + mp(f)[u(v+a) - u(v)]$ and $I(u(w)+\lambda\psi) = u(w) + mp(g)[u(v+a) - u(v)]$.

Since statement 2 is $mp(g) \geq mp(f)$, it is henceforth equivalent to $I(u(w) + \lambda\psi) - u(w) \geq \lambda(I(u(v) + \psi) - u(v))$, which is statement 3. $\square$

Proof of Theorem 10

Proof. The proof is omitted because it is basically the same as the proof of Theorem 9. $\square$

8.2 Other Materials on the Smooth Ambiguity Model

8.2.1 Example: Ambiguity and Ambiguity Attitude Compounded

Here is an example, under the smooth ambiguity models $\{(\phi^i, u^i, \mu^i)\}_{i=1,2}$, where DM 1 with a less concave ϕ^1 can be more ambiguity averse because μ^1 is more "dispersed". Suppose $u^1(\overline{m}_1) = u^2(\overline{m}_2) = 1$, $u^1(\underline{m}_1) = u^2(\underline{m}_2) = 0$, and $\phi^1(x) = -exp(-x)$ and $\phi^2(x) = -exp(-2x)$ for all $x \in [0,1]$. Notice that ϕ^1 is less concave because $\phi^2 = -(\phi^1)^2$ and the transformation $-x^2$ is concave for all $x \in \phi^1([0,1])$. Further suppose there are only two states $\{s_0, s_1\}$, μ^1 put $\frac{1}{2}$ on the point mass $p_0 = \delta_{s_0}$ and $\frac{1}{2}$ on the point mass $p_1 = \delta_{s_1}$, and μ^2 put $\frac{1}{2}$ on $p_{\frac{1}{2}-\epsilon}$, according to which s_1 happens with probability $\frac{1}{2} - \epsilon$, and $\frac{1}{2}$ on $p_{\frac{1}{2}+\epsilon}$, according to which s_1 happens with probability $\frac{1}{2} + \epsilon$. We want to find ϵ small enough so that DM 2, who has a more concave ϕ^2, is actually less ambiguity averse. In other words, the more concave ϕ^2 is over-compensated by the much more concentrated μ^2.

When p_0 and p_1 evaluate an utility profile $u^1(f)$ (or $u^2(g)$, for aligned f and g) at a and b, respectively, $p_{\frac{1}{2}-\epsilon}$ and $p_{\frac{1}{2}+\epsilon}$ evaluates it at $\frac{a+b}{2} - \epsilon(b-a)$ and $\frac{a+b}{2} + \epsilon(b-a)$, respectively. It can be shown that DM 1 is more ambiguity averse if and only if for all $a, b \in [0,1]$,

$$\left(\frac{1}{2}exp(-a) + \frac{1}{2}exp(-b)\right)^2 \geq \frac{1}{2}exp\big(-(a+b) - 2\epsilon(b-a)\big) + \frac{1}{2}exp\big(-(a+b) + 2\epsilon(b-a)\big).$$

Simplification yields that DM 1 is more ambiguity averse if and only if for all $\delta \in [0,1]$,

$$\frac{1}{2} + \frac{1}{4}exp(\delta) + \frac{1}{4}exp(-\delta) \geq \frac{1}{2}exp(2\epsilon\delta) + \frac{1}{2}exp(-2\epsilon\delta).$$

Take $\epsilon = \frac{1}{100}$, and then the inequality is satisfied, with strict inequality for all $\delta \in (0,1]$. This is because both the left hand side and the right hand side are 1 when $\delta = 0$, yet the left hand side increases much faster at any $\delta \in (0,1)$.

8.2.2 Result with Enriched Preferences

For this subsection, suppose, for $i \in \{1,2\}$, $\succcurlyeq^i$ is over second order acts, which are simple functions from Δ to $\mathcal{X}$ measurable to the Borel σ-algebra Σ_Δ of Δ (which is equipped with the weak* topology). We show that the characterization result works without Assumption CS.

Theorem 17. *Given smooth ambiguity models $\{(\phi^i, u^i, \mu^i)\}_{i=1,2}$ with the normalization that $u^1(\overline{m}_1) = u^2(\overline{m}_2)$ and $u^1(\underline{m}_1) = u^2(\underline{m}_2)$, under Assumption B, and the existence of an interval on which ϕ^1 and ϕ^2 are continuously differentiable, the following two statements are equivalent.*

1. *DM 1 is more ambiguity averse* than DM 2.*
2. *$\phi^1 = \psi \circ \phi^2$ for some concave function $\psi : \phi^2(u^2(\mathcal{X})) \to \mathbb{R}$, and, $\mu^1 = \mu^2$.*

Proof. (1) $\implies$ (2) is obvious and thus omitted.

For (1) $\impliedby$ (2), the proof takes the same roadmap as the proof of Theorem 6.

In step 1, we turn to the part $\mu^1 = \mu^2$. If $\mu^1 \neq \mu^2$, there is a simple function $h \in B_0(\Sigma_\Delta, \mathbb{R})$ such that $\int_\Delta E_p h d\mu^1(p) > \int_\Delta E_p h d\mu^2(p)$. By translation and scaling, and also because now $\succcurlyeq^i$ is over second order acts, we can without loss of generality have h being the utility profile of some aligned acts $f \in \mathcal{F}_{\overline{m}_1, \underline{m}_1}$ and $g \in \mathcal{F}_{\overline{m}_2, \underline{m}_2}$. Then the proof proceeds just like the step 1 of the proof to Theorem 6, which shows that statement (1) must be violated.

In step 2, we turn to the part $\phi^1 = \psi \circ \phi^2$ for some concave function $\psi : \phi^2(u^2(\mathcal{X})) \to \mathbb{R}$. Here we prove the result using only first order acts, and therefore this also delivers the characterization result when $\succcurlyeq^i$ are defined over first order acts (as in the main text) with $\mu^1 = \mu^2$ being directly assumed. If $\mu^1 = \mu^2$ has a support that is a singleton, we take ϕ^1, ϕ^2, and ψ to be the identity function. If $\mu^1 = \mu^2$ has a nontrivial support, take two distinct points q_1 and q_2 in that support and take a simple function $l \in B_0(\Sigma_{\mathcal{S}}, \mathbb{R})$ such that $E_{q_1} l > E_{q_2} l$. By translation and scaling, we can without loss of generality assume $\overline{l} = max_{s \in \mathcal{S}} l(s) = 1$ and $\underline{l} = min_s l(s) = 0$. The two open sets $\{p \in \Delta : E_p l > \frac{2}{3} E_{q_1} l + \frac{1}{3} E_{q_2} l\}$ and $\{p \in \Delta : E_p l < \frac{1}{3} E_{q_1} l + \frac{2}{3} E_{q_2} l\}$ must be assigned positive probability by μ, and therefore μ and l induces a non-degenerate probability distribution ρ on $[0,1]$ with $\rho(E) = \mu(\{p \in \Delta : E_p l \in E\})$ for all $E \in \Sigma_{\mathcal{S}}$. For any $a < b$, $\rho_{a,b}$ denotes the probability measure on $([a,b], \Sigma_{[a,b]})$ such that $\rho_{a,b}(E) = \rho(\{\frac{x-a}{b-a} : x \in E\})$ for all $E \in \Sigma_{[a,b]}$.

For any $a \in (0,1)$ and any $\lambda \in (0, 1-a)$, the utility profile $a + \lambda \cdot l \in B_0(\Sigma, [0,1])$ is the utility profile of some aligned acts $f \in \mathcal{F}_{\overline{m}_1, \underline{m}_1}$ and $g \in \mathcal{F}_{\overline{m}_2, \underline{m}_2}$. For this pair of aligned f and g statement (1) implies

$$(\phi^1)^{-1}\big[E_\mu \phi^1(\alpha + \lambda E_p l)\big] \leq (\phi^2)^{-1}\big[E_\mu \phi^2(\alpha + \lambda E_p l)\big].$$

For all $0 \leq a < b \leq 1$, by taking the above $\lambda = b - a$, we have

$$E_{\rho_{a,b}} \psi \circ \phi^2(x) \leq \psi\big(E_{\rho_{a,b}} \phi^2(x)\big),$$

which can be reformulated as for all $a, b \in \phi^2([0,1])$ and $a < b$,

$$E_{\rho'_{a,b}}\psi(x) \le \psi\big(E_{\rho'_{a,b}}x\big), \tag{8.4}$$

where $\rho'_{a,b}(E) = \rho_{(\phi^2)^{-1}(a),(\phi^2)^{-1}(b)}((\phi^2)^{-1}(E))$ for all $E \in \Sigma_{[a,b]}$.

We claim inequality (8.4) implies that ψ is concave on $\phi^2([0,1])$. Suppose $\psi(\alpha y + (1-\alpha)x) < \alpha\psi(y) + (1-\alpha)\psi(x)$ for some $\alpha \in (0,1)$ and $x < y \in \phi^2([0,1])$. Consider the line segment l_{xy} connecting $(x, \psi(x))$ to $(y, \psi(y))$, and, because ψ is continuous, take $a < \alpha y + (1-\alpha)x$ to be the largest value such that $(a, \psi(a))$ is on l_{xy} and $b > \alpha y + (1-\alpha)x$ to be the smallest value such that $(b, \psi(b))$ is on l_{xy} . Then, ψ restricted on (a, b) must be strictly below l_{xy} restricted on (a, b). Let $c = min\{argmin_{z\in[a,b]}\psi(z) - l_{xy}(z)\}$ and find $[e, f] \subset [a, b]$ such that $E_{\rho'_{e,f}}x = c$, which can be done because ϕ^2 is continuous and strictly increasing. Then, because the probability mass of ρ is not centered around a single point, we have $E_{\rho'_{e,f}}\psi(x) > \psi(E_{\rho'_{e,f}}x)$, a contradiction. □

8.2.3 Relaxation of Common Second-Order Belief

Baillon, Driesen, and Wakker (2012, hereafter BDW) present a method to compare ambiguity attitudes in the smooth ambiguity model without the common second-order belief assumption (common utility is still needed). Here we show that incorporating their idea into ours enables both taste and belief to vary across the DMs. We start with a much weaker assumption.[27]

Assumption TS (Twin Support) The support of μ^i consists of two mutually singular first-order beliefs $\{p^i, q^i\} \subset \Delta$.

We adapt BDW's technique of deriving preference midpoints to lotteries. Notice that by Assumption TS, there is an event $E^i \subset \mathcal{S}$ such that $p^i(E^i) = q^i((E^i)^c) = 1$. To streamline notation, for all $\alpha, \beta \in [0,1]$, $\big(\alpha, E^i; \beta, (E^i)^c\big) \in \mathcal{F}_{(\overline{x}_i, \underline{x}_i)}$ denotes the act for which $(\alpha\overline{m}_i + (1-\alpha)\underline{m}_i)$ is obtained on E^i and $(\beta\overline{m}_i + (1-\beta)\underline{m}_i)$ is obtained on $(E^i)^c$.

Definition. For any $\alpha_1, \alpha_2, \alpha_3 \in [0,1]$, α_2 is a preference midpoint between α_1 and α_3 for DM $i \in \{1,2\}$, denoted as $\alpha_1\alpha_2 \backsim^i \alpha_2\alpha_3$, if for some $\beta \in [0,1]$ and $F^i \in \{E^i, (E^i)^c\}$ where $p^i(E^i) = q^i((E^i)^c) = 1$, we have the following four acts that belong to $\mathcal{F}_{(\overline{m}_i, \underline{m}_i)}$ and satisfy

$$\big(\alpha_2, F^i; 0, (F^i)^c\big) \sim^i \big(\alpha_1, F^i; \beta, (F^i)^c\big),$$
$$\big(\alpha_3, F^i; 0, (F^i)^c\big) \sim^i \big(\alpha_2, F^i; \beta, (F^i)^c\big).$$

[27]This assumption was used by KMM to deliver what is assumed in the richness assumption 3.2 of BDW.

Following BDW, we can then formulate comparative ambiguity attitudes using preference midpoints.[28]

Definition. DM 1 is more ambiguity^s averse than DM 2 if for all $\alpha_1, \alpha_2^1, \alpha_2^2, \alpha_3 \in [0,1]$ such that $\alpha_1\alpha_2^i \backsim^i \alpha_2^i\alpha_3$ we have $\alpha_2^1 \le \alpha_2^2$.

We choose not to use this formulation in the main text because it does not have the flavor of tradeoffs between an ambiguous act and an ambiguity-free act. However, it does free the DMs from both common belief and common taste, as the following result shows.

Proposition 18. *Given Assumption B, smooth ambiguity models $\{(\phi^i, u^i, \mu)\}_{i=1,2}$ that satisfy Assumption TS, and the normalization $u^1(\overline{m}_1) = u^2(\overline{m}_2)$ and $u^1(\underline{m}_1) = u^2(\underline{m}_2)$, the following two statements are equivalent.*

1. *DM 1 is more ambiguity^s averse than DM 2.*
2. *ϕ^1 is more concave than ϕ^2.*

Proof. Given the normalization $u^1(\overline{m}_1) = u^2(\overline{m}_2) = 1$ and $u^1(\underline{m}_1) = u^2(\underline{m}_2) = 0$, acts of the form $\big(\alpha, E^i; \beta, (E^i)^c\big) \in \mathcal{F}_{(\overline{m}_i, \underline{m}_i)}$ are evaluated as $\mu^i(p^i)\phi^i(\alpha) + (1-\mu^i(p^i))\phi^i(\beta)$. Then, it is straightforward to verify that $\alpha_1\alpha_2 \backsim^i \alpha_2\alpha_3$ if and only if $\phi^i(\alpha_2) = \frac{1}{2}\phi^i(\alpha_1) + \frac{1}{2}\phi^i(\alpha_3)$.

Let $\psi = \phi^1 \circ (\phi^2)^{-1}$ and then we have (1) if and only if for all $\alpha_1, \alpha_3 \in [0,1]$ we have

$$(\phi^1)^{-1}\big[\frac{1}{2}\phi^1(\alpha) + \frac{1}{2}\phi^1(\beta)\big] \le (\phi^2)^{-1}\big[\frac{1}{2}\phi^2(\alpha) + \frac{1}{2}\phi^2(\beta)\big],$$

which is equivalent to

$$\frac{1}{2}\psi \circ \phi^2(\alpha) + \frac{1}{2}\psi \circ \phi^2(\beta) \le \psi\big[\frac{1}{2}\phi^2(\alpha) + \frac{1}{2}\phi^2(\beta)\big].$$

This is further equivalent to for all $\forall a, b \in \phi^2([0,1])$,

$$\frac{1}{2}\psi(a) + \frac{1}{2}\psi(b) \le \psi\big[\frac{1}{2}a + \frac{1}{2}b\big],$$

which is true if and only if (2) is true. □

[28]We do not use the word "probability midpoints" because these midpoints depend on the prizes $\{\overline{m}_i, \underline{m}_i\}$ used in the definition.

References

Abdellaoui, M. (2000): "Parameter-free elicitation of utility and probability weighting functions," *Management science*, 46, 1497–1512. 34

——— (2002): "A genuine rank-dependent generalization of the Von Neumann-Morgenstern expected utility theorem," *Econometrica*, 70, 717–736. 8

Abdellaoui, M. and P. Wakker (2005): "The likelihood method for decision under uncertainty," *Theory and decision*, 58, 3–76. 8

Akaike, H. (1974): "A new look at the statistical model identification," *IEEE Transactions on Automatic Control*, 19, 716–723. 32

Allais, M. (1953): "L'extension des théories de l'équilibre économique général et du rendement social au cas du risque," *Econometrica, Journal of the Econometric Society*, 269–290. 33

Alon, S. and G. Gayer (2016): "Utilitarian preferences with multiple priors," *Econometrica*, 84, 1181–1201. 15

Alon, S. and D. Schmeidler (2014): "Purely subjective maxmin expected utility," *Journal of Economic Theory*, 152, 382–412. 36

Anscombe, F. J. and R. J. Aumann (1963): "A definition of subjective probability," *Annals of mathematical statistics*, 34, 199–205. 2, 5, 9

Arrow, K. J. (1965): *Aspects of the theory of risk-bearing*, Yrjö Jahnssonin Säätiö. 2

——— (1971): *Essays in the theory of risk-bearing*, Markham Publishing Company,Chicago. 20

Baillon, A., H. Bleichrodt, C. Li, and P. P. Wakker (2019): "Belief Hedges: Applying ambiguity measurements to all events and all ambiguity models," *working paper*. 8, 10, 12

Baillon, A., B. Driesen, and P. P. Wakker (2012): "Relative concave utility for risk and ambiguity," *Games and Economic Behavior*, 75, 481–489. 16, 44

Baillon, A., Z. Huang, A. Selim, and P. P. Wakker (2018): "Measuring ambiguity attitudes for all (natural) events," *Econometrica*, 86, 1839–1858. 4, 8, 12

Baillon, A. and L. Placido (2019): "Testing constant absolute and relative ambiguity aversion," *Journal of Economic Theory*, 181, 309–332. 8, 20, 23

Basu, P. and F. Echenique (2018): "Learnability and Models of Decision Making under Uncertainty." *working paper*. 28

Bewley, T. F. (2002): "Knightian decision theory. Part I," *Decisions in economics and finance*, 25, 79–110. 13

Cerreia-Vioglio, S., F. Maccheroni, and M. Marinacci (2019): "Ambiguity Aversion and Wealth Effects," *forthcoming at Journal of Economic Theory*. 2, 7, 8, 20, 35

Cerreia-Vioglio, S., F. Maccheroni, M. Marinacci, and L. Montrucchio (2011a): "Complete monotone quasiconcave duality," *Mathematics of Operations Research*, 36, 321–339. 32

——— (2011b): "Uncertainty averse preferences," *Journal of Economic Theory*, 146, 1275–1330. 2, 31

Chateauneuf, A. and J. H. Faro (2009): "Ambiguity through confidence functions," *Journal of Mathematical Economics*, 45, 535–558. 2, 30

Cherbonnier, F. and C. Gollier (2015): "Decreasing aversion under ambiguity," *Journal of Economic Theory*, 157, 606–623. 20

Deakin, J., M. Aitken, T. Robbins, and B. J. Sahakian (2004): "Risk taking during decision-making in normal volunteers changes with age," *Journal of the International Neuropsychological Society*, 10, 590–598. 3

Dimmock, S. G., R. Kouwenberg, and P. P. Wakker (2016): "Ambiguity attitudes in a large representative sample," *Management Science*, 62, 1363–1380. 4, 8, 33

Dohmen, T., A. Falk, D. Huffman, U. Sunde, J. Schupp, and G. G. Wagner (2011): "Individual risk attitudes: Measurement, determinants, and behavioral consequences," *Journal of the European Economic Association*, 9, 522–550. 3

Ellsberg, D. (1961): "Risk, ambiguity, and the Savage axioms," *The quarterly journal of economics*, 643–669. 4

Epstein, L. G. (1999): "A Definition of Uncertainty Aversion," *Review of Economic Studies*, 66, 579–608. 2, 4, 8, 11, 28, 35, 36

Epstein, L. G. and J. Zhang (2001): "Subjective probabilities on subjectively unambiguous events," *Econometrica*, 69, 265–306. 2, 35, 36

Gajdos, T., T. Hayashi, J.-M. Tallon, and J.-C. Vergnaud (2008): "Attitude toward imprecise information," *Journal of Economic Theory*, 140, 27–65. 8

Ghirardato, P., F. Maccheroni, and M. Marinacci (2004): "Differentiating ambiguity and ambiguity attitude," *Journal of Economic Theory*, 118, 133–173. 29

Ghirardato, P. and M. Marinacci (2002): "Ambiguity made precise: A comparative foundation," *Journal of Economic Theory*, 102, 251–289. 2, 4, 6, 7, 8, 11, 22

Gilboa, I. (1987): "Expected utility with purely subjective non-additive probabilities," *Journal of mathematical Economics*, 16, 65–88. 8

Gilboa, I., F. Maccheroni, M. Marinacci, and D. Schmeidler (2010): "Objective and subjective rationality in a multiple prior model," *Econometrica*, 78, 755–770. 13

Gilboa, I. and D. Schmeidler (1989): "Maxmin expected utility with non-unique prior," *Journal of mathematical economics*, 18, 141–153. 1, 6, 12

Gonzalez, R. and G. Wu (1999): "On the shape of the probability weighting function," *Cognitive psychology*, 38, 129–166. 33

Grant, S. and B. Polak (2013): "Mean-dispersion preferences and constant absolute uncertainty aversion," *Journal of Economic Theory*, 148, 1361–1398. 20

Hansen, L. P. and T. J. Sargent (2008): *Robustness*, Princeton university press. 33

Harsanyi, J. C. (1955): "Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility," *Journal of political economy*, 63, 309–321. 11

Hartmann, L. (2018): "A Definition of Perceived Ambiguity," *working paper*. 30

Hill, B. (2019): "A Non-Bayesian Theory of State-Dependent Utility," *forthcoming at Econometrica*. 8

Kihlstrom, R. E., A. E. Roth, and D. Schmeidler (1981): "Risk aversion and solutions to Nash's bargaining problem," in *Game Theory and Mathematical Economics, ed. by O. Moeschlin and D. Pallaschke*, Amsterdam: North-Holland. 2

Klibanoff, P., M. Marinacci, and S. Mukerji (2005): "A Smooth Model of Decision Making under Ambiguity," *Econometrica*, 73, 1849–1892. 1, 2, 9, 14

Knight, F. H. (1921): *Risk, Uncertainty and Profit*, Boston, New York, Houghton Mifflin Company. 2

Kopylov, I. (2016): "Subjective probability, confidence, and Bayesian updating," *Economic Theory*, 62, 635–658. 8

Lang, M. (2017): "First-Order and Second-Order Ambiguity Aversion," *Management Science*, 1254–1269. 35

Maccheroni, F., M. Marinacci, and D. Ruffino (2013): "Alpha as Ambiguity: Robust Mean-Variance Portfolio Analysis," *Econometrica*, 81, 1075–1113. 26

Maccheroni, F., M. Marinacci, and A. Rustichini (2006): "Ambiguity Aversion, Robustness, and the Variational Representation of Preferences," *Econometrica*, 74, 1447–1498. 2, 5, 9, 32

Machina, M. J. and D. Schmeidler (1992): "A more robust definition of subjective probability," *Econometrica: Journal of the Econometric Society*, 745–780. 35

Pratt, J. W. (1964): "Risk Aversion in the Small and in the Large," *Econometrica*, 32, 122–136. 2

Quiggin, J. (1982): "A theory of anticipated utility," *Journal of Economic Behavior & Organization*, 3, 323–343. 34

Ross, S. A. (1981): "Some stronger measures of risk aversion in the small and the large with applications," *Econometrica: Journal of the Econometric Society*, 621–638. 2

Savage, L. J. (1954): *The foundations of statistics*, New York: John Wiley and Sons. (Second addition in 1972, Dover). 5

Schmeidler, D. (1989): "Subjective Probability and Expected Utility without Additivity," *Econometrica*, 57, 571–87. 8, 28

Strzalecki, T. (2011): "Axiomatic Foundations of Multiplier Preferences," *Econometrica*, 79, 47–73. 33

Wakker, P. (1989): *Additive Representations of Preferences: A New Foundation of Decision Analysis*, Dordorecht. 36

Wang, F. (2019): "A Dialogue Between Risk and Ambiguity: Ambiguity Aversion with Rank-Dependent Utility," *working paper*. 8, 18, 34

Xue, J. (2018): "Preferences with changing ambiguity aversion," *Economic Theory*, 1–60. 20

Yaari, M. E. (1969): "Some remarks on measures of risk aversion and on their uses," *Journal of Economic Theory*, 1, 315–329. 2