Belief hedges: Measuring ambiguity for all events and all models [☆]

Aurélien Baillon, Han Bleichrodt, Chen Li, Peter P. Wakker [*]

Erasmus School of Economics, Erasmus University Rotterdam, Rotterdam, the Netherlands

Abstract

We introduce belief hedges, i.e., sets of events whose uncertain subjective beliefs neutralize each other. Belief hedges allow us to measure ambiguity attitudes without knowing those subjective beliefs. They lead to improved ambiguity indexes that are valid under all popular ambiguity theories. Our indexes can be applied to real-world problems and do not require expected utility for risk or commitments to two-stage optimization, thereby increasing their descriptive power. Belief hedges make ambiguity theories widely applicable.

1. Introduction

Hedging is a central concept in finance. It turns uncertain monetary outcomes ("gambles") into certainties without using any further information about the relevant uncertainties. A hedge combines a properly chosen set of gambles so that their uncertainties neutralize each other and

the risk-neutral value is obtained irrespective of what those uncertainties are. This paper introduces an analog for subjective beliefs, called belief hedges. A well-known problem in measuring ambiguity[1] attitudes is that there is uncertainty about an agent's subjective beliefs, which may confound the measurement. A belief hedge combines a properly chosen set of events, so that uncertain subjective beliefs neutralize each other. We can then calibrate ambiguity neutrality (formalized in Observation 20) irrespective of the uncertain beliefs, and measure ambiguity attitudes without needing any information about those beliefs.

Belief hedges can be directly applied to the real-life uncertainties that are relevant for applications. They do not need artificial events such as Ellsberg urns or experimenter-specified probability intervals, which increases external validity and the motivation of clients and subjects. Using our belief hedges, we introduce general indexes of ambiguity aversion and insensitivity.

Baillon et al. (2018) were the first to measure ambiguity attitudes without needing information about subjective beliefs. We show that their domain of events is the special case of belief hedges for three-fold partitions. They did not justify their method theoretically. We will do so for our generalization and, consequently, we will also justify their method. Our extension beyond three-fold partitions gives the desired flexibility for applications (Examples 4, 14, and 22). Our main contribution is that we introduce the relevant general concepts: belief hedges and the indexes of ambiguity aversion and insensitivity that we derive from them. We give preference foundations for these concepts, supporting their validity. We also show that these concepts are based on underlying econometric principles. Section 8 gives further details on our contribution to Baillon et al. (2018).

Our indexes generalize most ambiguity indexes that have been proposed in the literature. They are thus valid under many ambiguity theories. They do not need expected utility for risk, or two-stage stimuli and dynamic decision principles, making them descriptively valid and tractable. Our indexes can directly be elicited from preferences and do not require data fitting or additional assumptions about parametric specifications or the underlying error model. They show, in particular, how many indexes in the literature can be made directly observable.

A detailed outline is as follows. The first part of our paper gives a model-free introduction of belief hedges and our indexes of ambiguity attitudes. Following basic definitions (§§2.1–2.3), §2.4 introduces belief hedges for ambiguity aversion and a corresponding aversion index—an average ambiguity premium. This section, while elementary, conveys the main novelty of belief hedges, explaining why they make artificial ambiguities (e.g., Ellsberg urns) redundant. Section 2.5 presents our second index of ambiguity, which captures insensitivity to changes in likelihood.

We interpret aversion as a motivational component of ambiguity attitude and insensitivity as a cognitive component. In support of this interpretation, §2.7 shows formally that the indexes are orthogonal and, therefore, capture distinct components in the variance of the data. Section 3 gives examples of belief hedges, illustrates their tractability, and gives a preference foundation of our indexes. It shows how belief hedges can be used to handle complex empirical problems.

The second part of this paper, starting in §4, is theoretical. We extend our indexes to many outcomes and first consider models, such as the smooth model, in which our indexes can be

---

[1] Ambiguity refers to uncertain events for which no probabilities are known. Risk refers to the case of known probabilities.

outcome dependent. We then move on to outcome independence, which holds in many models including biseparable utility, rank-dependent/Choquet expected utility, prospect theory, and the various multiple prior models. The notation of the first part of this paper, which did not express outcome dependence, can then still be used and is most convenient.

Section 7 shows that our indexes generalize and unify most indexes proposed before. They also agree with the common qualitative orderings of ambiguity attitudes, such as being more ambiguity averse. Consequently, the arguments that have been put forward in the literature to support existing indexes and orderings give broad theoretical support to our indexes. The main text ends with discussions and a conclusion (§§8-9). Proofs are in the appendix.

2. Belief hedges

This section defines belief hedges and provides theoretical justifications.

2.1. Basic definitions

S denotes a *state space*, finite or infinite. Its subsets are *events*. X denotes a set of *outcomes*, again, finite or infinite. Outcomes can be money amounts, health states, and so on. *Acts* map S to X and are finite-valued.[2] Act $\gamma_E \beta$ assigns outcome γ to event E and outcome β to all other states. We further assume that *lotteries* $\gamma_p \beta$ (receiving outcome γ with probability p and β with probability $1 - p$) are available. We use the term *prospect* for both acts and lotteries. A *preference relation* $\succcurlyeq$ is given over prospects, with the usual notation $\preccurlyeq, \succ, \prec$, and $\sim$. We assume weak ordering throughout (completeness and transitivity). As usual, we identify constant acts and degenerate lotteries with outcomes. This implies $\gamma = \gamma_S \beta = \gamma_1 \beta$, and $\succcurlyeq$ now also applies to outcomes. An event is *null* if its outcomes never affect preference. *Monotonicity* means: (i) weakly improving an outcome of a prospect weakly improves the prospect; (ii) strictly improving an outcome of an act in a nonnull event strictly improves the act; (iii) strictly improving an outcome of a lottery with positive probability strictly improves the lottery.

A *measurement design* $\mathcal{H}$ is a finite collection of events. It describes the events that will be used to define and measure the ambiguity indexes. A central question in our analysis will be which designs to use for these purposes. In most of this paper (except §3) $\mathcal{H}$ is fixed, and then dependencies on it can be dropped from the notation. By $\{E_1, \ldots, E_n\}$, the *design atoms*, or *atoms* for short, we denote the smallest nonempty intersections of events in $\mathcal{H}$. Belief hedging will imply that the atoms partition S, covering all states. The E_js are the "atoms" of the smallest (finite) algebra of events generated by $\mathcal{H}$. For $E \in \mathcal{H}$, $|E|$ denotes the number of atoms contained in E. The Greek nu (ν) denotes the *normalized event size*, or *event size* for short, with $\nu(E) = \frac{|E|}{n}$. Thus, $\nu(S) = 1$. Section 3 shows that under empirically plausible assumptions, our indexes are largely independent of the design (and, thus, of ν).

Throughout this paper, statistics refer to $\mathcal{H}$, defined as usual. For functions $F, G : \mathcal{H} \to \mathfrak{R}$, by $\overline{F} = \frac{\sum_{E \in \mathcal{H}} F(E)}{|\mathcal{H}|}$ we denote the average of F; $Var(F)$ denotes F's variance; $Cov(F, G)$ denotes the covariance of F and G. Details are in Appendix A. We define the *sensitivity* of F with respect to G as $\frac{Cov(F,G)}{Var(G)}$. It is a first-order approximation of how much F will change on average if G changes by one unit.

---

[2] We can endow S with an algebra of events containing all singletons $\{s\}$, and consider only measurable acts. Then nothing in this paper changes.

2.2. Structural assumption

To achieve maximal generality and maximal accessibility, Sections 2 and 3 introduce and analyze our concepts under minimal assumptions, which hold for all ambiguity models. We relax these assumptions in later sections. Sections 2 and 3 assume:

Assumption 1 *(Two outcomes).* $X = \{\gamma, \theta\}$, with $\gamma \succ \theta$. □

Later sections consider multiple outcomes. For a large class of models (§7), our concepts and results are independent of outcomes, and Assumption 1 is not restrictive for them. We assume that a *matching probability* $m(E)$ exists for every event E, defined by

$$\gamma_E \theta \sim \gamma_{m(E)} \theta. \tag{1}$$

Monotonicity implies that m is unique. Dimmock et al. (2016, Theorem 3.1) and Gul and Pesendorfer (2020, "risk equivalent") showed that matching probabilities are well suited to analyze ambiguity attitudes because, under many ambiguity models, they capture everything relevant to ambiguity attitudes. There is no need to measure risk attitudes, utilities, probability weighting, and so on. Ambiguity reflects how $m(\cdot)$ deviates from a probability measure and, thus, from additivity. For example, ambiguity aversion implies $m(E) + m(E^c) < 1$, violating additivity.

We summarize the structural assumptions for the entire paper:

Assumption 2 *(Structural assumption).* $\succcurlyeq$ is a monotonic weak order over acts (finite-valued measurable maps from S to X) and lotteries (two-valued probability distributions over X). Each event has a matching probability. □

2.3. Ambiguity neutrality

Definition 3. $\succcurlyeq$ is *ambiguity neutral* if m is a probability measure. □

Ambiguity neutrality means that subjective beliefs are treated the same way as objective beliefs. We define it for two outcomes here, but Observation 20 will show that general ambiguity neutrality is implied under many ambiguity models. Ambiguity neutrality is violated in the following example. We chose it because it refers to an existing experiment, showing the practical relevance of our approach. As a price to pay, it involves some game-theoretic details.

Example 4 *(Running example).* The agent is player 1 in the following two-player minimum effort coordination game from Goeree and Holt (2001), in the version analyzed theoretically by Eichberger and Kelsey (2011). There are six effort levels $\{115, 125, 135, 145, 155, 165\}$. They are acts, $f_1, \ldots, f_6$ (with "f" referring to first player) and the agent has to choose one of them. Player 2 also has to choose one of these six effort levels simultaneously and independently. For the agent, these are states, denoted $s_1, \ldots, s_6$ (with "s" referring to second player) and they are a source of uncertainty.

The players receive the outcome $\min\{f, s\} - c \times e$, where f is chosen by player 1, s by player 2, e denotes own effort level, and $c = 0.9$ denotes the marginal cost. The best outcome possible for both players results from the joint maximum effort $f_6 = s_6 = 165$, constituting the most favorable equilibrium. However, if one player chooses a lower effort level, e.g., player 2 chooses s_3, then the optimal choice for the other player is the same, f_3. More effort then leads to a loss. A

less favorable equilibrium results. In this sense, a choice f_6 is risky, gambling on trust and cooperation, and f_1 is safe. In experiments, most players choose the safe, low effort (Table 1 below).

In this example, we assume that we can also observe the agent's matching probabilities through the following indifferences[3]:

$$15_{s_1}0 \sim 15_{0.50}0; \text{ i.e., } m(s_1) = 0.50;$$
$$15_{s_2}0 \sim 15_{s_3}0 \sim 15_{s_4}0 \sim 15_{s_5}0 \sim 15_{0.20}0; \text{ i.e., } m(s_2) = \cdots = m(s_5) = 0.20;$$
$$15_{s_6}0 \sim 15_{0.30}0; \text{ i.e., } m(s_6) = 0.30;$$
$$0_{s_1}15 \sim 0_{0.55}15; \text{ i.e., } m(s_1^c) = 0.45;$$
$$0_{s_2}15 \sim 0_{s_3}15 \sim 0_{s_4}15 \sim 0_{s_5}15 \sim 0_{0.30}15; \text{ i.e., } m(s_2^c) = \cdots = m(s_5^c) = 0.70;$$
$$0_{s_6}15 \sim 0_{0.35}15; \text{ i.e., } m(s_6^c) = 0.65.$$

The agent violates ambiguity neutrality: m is not a probability measure because it violates additivity. For instance, $m(s_1) + m(s_1^c) = 0.95 < 1$.

We will use this example to illustrate the techniques of our paper, and will summarize and discuss the results in Example 22. □

2.4. Belief hedging for ambiguity aversion

To measure ambiguity aversion, several papers used differences $P(E) - m(E)$, where P denotes subjective probabilities reflecting ambiguity-neutral beliefs, called *a-neutral probabilities*.[4] These differences reflect an ambiguity premium, i.e., a willingness to pay—in probability (belief) units—to avoid ambiguity. This premium increases with the degree of ambiguity aversion (Dimmock et al., 2016; Kahn and Sarin, 1988; Viscusi and Magat, 1992). Ideally, with observations $P(E) - m(E)$ available for a number of events E, we would like to define our aversion index as the average

$$\overline{P - m}. \tag{2}$$

The problem with this definition is that P is usually unknown. For Ellsberg urns, we can derive P from symmetry assumptions, but for natural events this cannot be done. In Example 4, P need not agree with the actual choice percentages in Table 1, as those are unknown to the agent.

Our solution is simple. We ensure, through Definition 5 below, a fixed and known average level of P:

$$\overline{P} = \frac{1}{2} \quad \text{for all } P. \tag{3}$$

Table 1
Choice percentages in Goeree and Holt (2001).

$s_1 = 115$	$s_2 = 125$	$s_3 = 135$	$s_4 = 145$	$s_5 = 155$	$s_6 = 165$
50	18	5	7	5	15

---

[3] Such belief measurements are commonly incentivized by means of a random incentive system that enhances isolation, avoiding income effects, hedging effects, and interactions between the game played and the belief measurement.

[4] They can be interpreted as the beliefs of the ambiguity neutral twin of the agent, i.e., the beliefs if the agent changed into ambiguity neutral but in all other respects remained the same.

To prepare for the concept relevant here (belief hedges), we present a condition that is not only sufficient, but also necessary, for Eq. (3):

Definition 5. $\mathcal{H}$ is *level-hedged*, or *l-hedged* for short, if:

$$\text{each state } s \text{ appears in exactly half of the events in } \mathcal{H}. \tag{4}$$

Equivalent is that each atom E_i appears in exactly half of the elements of $\mathcal{H}$ (can be seen via any $s \in E_i$). This implies $\overline{P} = 1/2$, first for all degenerate probability measures assigning probability 1 to E_i, and then for all their convex combinations, i.e., for all P. For applications, the most tractable special case is when $\mathcal{H}$ is complementation closed.[5] We multiply $\overline{P - m} = \frac{1}{2} - \overline{m}$ by 2 for normalization explained later:

Definition 6. If l-hedging (Eq. (4)) holds, then the *index of ambiguity aversion* is

$$b = 1 - 2\overline{m}. \tag{5}$$

Using l-hedging, we have captured Eq. (2) without needing to know P. As mentioned, our index reflects how much success probability one is willing to give up to avoid ambiguity. In the Anscombe-Aumann framework (expected utility for risk), it reflects the proportion of success-utility the agent is willing to pay. For linear utility, which is reasonable for small stakes, it is the proportion of the winning prize the agent is willing to give up to avoid ambiguity. In Example 4, we obtain $b = 0.08$, which suggests weak ambiguity aversion.

2.5. Belief hedging for insensitivity

Theoretically and normatively motivated ambiguity models have focused on ambiguity aversion, a motivational component of ambiguity attitude, and the topic of the preceding section.[6] However, empirical studies have found richer phenomena (Anantanasuwong et al., 2020; l'Haridon et al., 2018; Kocher et al., 2018). Whereas strong ambiguity aversion has indeed been reported for likely events, it gets weaker for events of moderate likelihood, and turns into ambiguity seeking for low likelihood events (Trautmann and van de Kuilen, 2015).[7]

The aforementioned pattern suggests a tendency to treat bets on events as fifty-fifty bets, with insufficient discriminatory power and insufficient responsiveness to changes in likelihoods in the middle region. It is similar to the insensitivity reflected by inverse-S probability weighting for risk, where weights in the middle are also moved toward fifty-fifty (Fehr-Duda and Epper, 2012; Wu and Gonzalez, 1999). We incorporate this insensitivity as a second, cognitive component of ambiguity attitude. It reflects a poor understanding of ambiguity. The agent takes ambiguous events (too much) as one blur. We use the term a(mbiguity-generated) insensitivity to refer to the insensitivity generated by ambiguity.

---

[5] However, this condition is not necessary and sufficient to serve in our axiomatization. For example, it is violated if S contains seven states and $\mathcal{H}$ contains all three- and six-state events. Then l-hedging still holds.

[6] An exception is Gul and Pesendorfer's (2015, p. 467, 471) Hurwicz expected utility, which explicitly allows for a-insensitivity (our term).

[7] These phenomena are reflected for losses, leading to a four-fold pattern. Overall, for losses there is more ambiguity seeking than aversion. Reflection can readily be accommodated by reflecting our parameters for losses or using dual functionals there. We focus on gains in this paper.

Because insensitivity reflects insufficient responsiveness of matching probabilities to changes in the a-neutral probabilities P, we would ideally like to base our insensitivity index on a measure of this responsiveness. The most common candidate is the sensitivity measure

$$\frac{Cov(m, P)}{Var(P)}. \tag{6}$$

This index has been widely used as the slope in regressions (Hill et al., 2008, Eq. (2.7)) and as β in the CAPM model in finance (Hull, 2017). It captures the average derivative of m with respect to P (in our domain of nonextreme events), i.e., the average change in m if P changes by one unit.

In the ε-contamination model, a tractable subclass of α-maxmin multiple prior models (defined later), our insensitivity index coincides with the size of the set of priors (§7.3). In general, the larger the set of priors (perception of ambiguity), the more events are treated alike, corresponding to lower discriminatory power. That is, there is more insensitivity. In the extreme case where the set of priors contains all priors, all nontrivial events E are treated alike, with all $\beta_E \alpha$ indifferent and with, indeed, maximal insensitivity. Section 7 provides similar results for other popular ambiguity theories, where insensitivity has often been interpreted as perception of ambiguity. Our insensitivity index shows a way to directly measure this based on revealed preferences.

To measure our insensitivity index, we again face the problem that the a-neutral P is unknown. Our solution is, again, to ensure that P does not matter. We can then replace any P by the event size ν, as if P were uniform over atoms. The idea is to ensure that the event size ν properly reflects the probability P (over events of the same size) in the sense that they co-vary perfectly with each other. The following condition is necessary and sufficient for this purpose. In the summation below, for fixed s we sum over all E.

Definition 7. $\mathcal{H}$ is ν (*ariation*)-*hedged* if:

$$\sum_{E \ni s} \nu(E) \quad \text{is the same for each fixed state } s. \tag{7}$$

This condition requires that the total size of events containing each fixed state s (equivalently, containing each atom E_i, through any $s \in E_i$), is a constant. It is satisfied in Example 4, where the sum of event sizes is 26 for each s. This condition is crucial in ensuring that the approximation in Eq. (8) below is proper. We provide an intuitive interpretation in the next subsection and present a full mathematical derivation in Appendix A.

To define our insensitivity index, we need one more structural assumption to avoid degeneracy.

Assumption 8 (*Nondegeneracy*). $\mathcal{H}$ does not contain $\emptyset$ or S. All atoms E_j are nonnull. Further, ν is not constant on $\mathcal{H}$. $\square$

Insensitivity concerns intermediate events away from the extremes and, therefore, the first part of the Assumption excludes the extreme events.[8] This entails no loss of information because the m values of $\emptyset$ and S are 0 and 1, respectively, by monotonicity. In the second part of the

---

[8] In the terminology of Wakker (2010, §7.7), we focus on the insensitivity region. Boundary restrictions can be used to define this region. Because our theorems are valid irrespective of what those regions are, we do not discuss them in this paper. For applications, we recommend using events in the measurement design with a-neutral probabilities between 0.05 and 0.95.

Assumption, null events do not affect preference and can therefore be removed from the atoms by joining them with a nonnull atom (with the obvious adaptation of $\mathcal{H}$). This second part further serves to stay away from extreme events. In the final part, event size needs to be nonconstant—also after excluding $\emptyset$ and S—because we derive insensitivity from variations in event size. This implies $n \geq 3$. The condition ensures that $Var(\nu)$ is positive, so that the ratios below are well-defined.

Definition 9. $\mathcal{H}$ is a *belief hedge* if both l-hedging and ν-hedging hold. □

The following definition is based on Eq. (6), with P replaced by ν.

Definition 10. If Assumption 8 holds and $\mathcal{H}$ is a belief hedge, then the *index of a(mbiguity-generated-)insensitivity* is

$$a = 1 - \frac{Cov(m, \nu)}{Var(\nu)}. \tag{8}$$

Because *Cov* and *Var* concern variation within $\mathcal{H}$, they can be calculated exactly from the matching probability data collected for all events in $\mathcal{H}$. Corollary 13 below gives a simple special case of Eq. (8) that can readily be calculated using paper and pencil. Whereas the aversion parameter captures how much success probability is lost due to ambiguity, the insensitivity index captures how much of the *changes* in probability is lost due to ambiguity (away from the extreme likelihoods). It thus captures the degree of underreaction to new information and is relevant, for example, in evaluations of precautionary measures. An index $a = 0.43$ (as in Example 4) means that the agent underestimates the marginal benefits of precautionary measures by a factor of almost 2.

Throughout the rest of the paper we assume Assumptions 2 and 8, explicitly in theorems and implicitly elsewhere. We end this subsection with a remark about the aversion index. The common empirical pattern of ambiguity seeking for unlikely events and (strong) ambiguity aversion for likely events implies that we would have underestimated [overestimated] the average $\overline{P - m}$ if we had included mainly unlikely [likely] events in our belief hedge. L-hedging avoids such biases by taking the average event-size equal to $1/2$.

2.6. Intuitive explanation of ν-hedging

To understand the intuition of ν-hedging, assume that it is violated (but l-hedging holds). More specifically, assume that there are states s and t such that the summation in Eq. (7) is smaller for s ("small") than for t. Because both states appear in half of the events in $\mathcal{H}$ by l-hedging, this means that t appears more often in big events in $\mathcal{H}$ than s, and, vice versa, that s appears more often in small events than t. Assume that we change a probability measure P by moving some probability mass from t to s. Then, within $\mathcal{H}$, the total probability of big events decreases, and that of small events increases. It is plausible that the total m value assigned to big [small] events then also decreases [increases]. That is, our index of a-insensitivity increases. However, this change is due to a change in the additive probability and not a change in ambiguity. Factors beyond ambiguity then confound our index. To avoid this, we impose ν-hedging.

In Example 4, ν-hedging holds. For all probability measures P, the sum of the probabilities of all small events of size $1/6$ is always 1, and the sum of the probabilities of all large events of size $5/6$ is always 5. Because the sums of the probabilities are the same for each P (including ν), the

extent to which the matching probabilities m over- or underweight events in $\mathcal{H}$ is independent of P. It must be due to deviations from probabilities P (ambiguity).

Because the sensitivity of m with respect to P is independent of P, we obtain the following approximation:

$$\frac{Cov(m, P)}{Var(P)} \approx \frac{Cov(m, \nu)}{Var(\nu)}. \tag{9}$$

This approximation justifies our index of a-insensitivity. Appendix A shows that Eq. (9) provides a good first-order approximation under common econometric assumptions. Proposition 24 in Appendix A shows some empirically plausible cases in which the equality is even exact.

2.7. Theoretical justifications of belief hedges

We first show formally that our indexes correctly classify ambiguity neutrality (and, accordingly, ambiguity aversion/seeking and (in)sensitivity) and that they have been properly normalized, allowing comparisons across studies. The following theorem shows that belief hedges are not only sufficient but also necessary for our purposes.

Theorem 11. *Under Assumptions 1, 2, and 8, $b = a = 0$ for all ambiguity neutral agents if and only if the measurement design $\mathcal{H}$ is a belief hedge. Then the supremum value of both b and a is 1.*[9] $\square$

Psychologically, we interpret aversion and insensitivity as two distinct components of ambiguity attitudes, motivational and cognitive. Theorem 12 shows that a mathematical orthogonality of the indexes supports this psychological interpretation. The proof of the theorem in Appendix B gives formalizations. Although conceptual orthogonality does not imply empirical orthogonality, Anantanasuwong et al. (2020) did find evidence for empirical orthogonality.

Theorem 12. *Under Assumptions 1, 2, and 8, the indexes a and b capture orthogonal components of the variance of the data.* $\square$

3. Which design to use and a preference foundation

This section examines variations in belief hedges, i.e., the measurement domains.

3.1. Examples of belief hedges

Baillon et al.'s (2018) experiment assumed three nonnull atoms $\{E_1, E_2, E_3\}$ and a full design $\mathcal{H} = \{E_1, E_2, E_3, E_1 \cup E_2, E_1 \cup E_3, E_2 \cup E_3\}$ denoted $\mathcal{H}(E_1, E_2, E_3)$. We write $\overline{m_s} = \frac{m(E_1) + m(E_2) + m(E_3)}{3}$ and $\overline{m_c} = \frac{m(E_1 \cup E_2) + m(E_1 \cup E_3) + m(E_2 \cup E_3)}{3}$. They used the following definitions, which by substitution (Appendix B) are identical to ours.

Corollary 13. $b = 1 - \overline{m_c} - \overline{m_s}$ *and* $a = 3(\frac{1}{3} - (\overline{m_c} - \overline{m_s}))$. $\square$

---

[9] Monotonicity excludes constancy of m. If we relax this condition, the supremum values can occur as maxima: $b = 1$ if m is constant 0 and $a = 1$ for any constant m.

We next give some other tractable examples of belief hedges. $\mathcal{H}$ is a belief hedge if for every $i < n$, every state (a) appears equally often in an event of size i; and (b) it does so with an overall frequency $\frac{1}{2}$. This includes all cases where l-hedging holds and $\mathcal{H}$ satisfies symmetry with respect to the atoms: for all $i \neq j$ and all $E_i, E_j \in \mathcal{H}$: if an event in $\mathcal{H}$ contains E_i but not E_j, then $\mathcal{H}$ also contains that event with E_i replaced by E_j. This is satisfied if $\mathcal{H}$ is the *full design*, i.e., it contains all unions of E_js except S and $\emptyset$, as in Baillon et al. (2018) with $\mathcal{H}\{E_1, E_2, E_3\}$. It is also satisfied if $\mathcal{H}$ is the *basic design*, i.e., it contains all one-atom events and their complements. Further, all disjoint unions of belief hedges are belief hedges.

The examples show that there is much flexibility in belief hedges. The smallest one possible is Baillon et al.'s (2018) design: the full design of a three-fold partition of S. Obviously, our indexes become more valid and reliable when $\mathcal{H}$ is richer. As default, we recommend a basic design with a partition of S that specifies all relevant uncertainties,[10] such as the six possible effort levels of the opponent in Example 4. This design involves all relevant atoms, considers both likely and unlikely events (where ambiguity is strongest), and grows linearly with the number of atoms so that it is tractable. An important advantage of the basic design (as well as richer designs) is that we get enough equalities to estimate the a-neutral probabilities as well (Example 22). In many real-world situations, control over the data received is limited, and general belief hedges give useful flexibility.

In the basic design of Example 4, we could, at will, add $\{\{s_1, s_2, s_3\}, \{s_4, s_5, s_6\}\}$, or any other pair of disjoint three-state events if these events are relevant or expected to give interesting behavior (see Example 14 below). We could further select events that subjects easily relate to or that avoid biases. We leave experimental implementations and real-world applications to future studies.

3.2. Non-uniform sources

In general, different designs do not need to give the same indexes, as the following example, also discussed by Machina (2011), shows.

Example 14. Consider an Ellsberg urn with 90 balls numbered 1-90, the first 30 red, the last 60 black or yellow in an unknown proportion. For E_1 (red), E_2 (non-red and odd), E_3 (non-red and even), the corresponding design $\mathcal{H}(E_1, E_2, E_3)$ will suggest ambiguity neutrality with $b = a = 0$. However, for E_1 (red), E_2 (black), E_3 (yellow), the corresponding design $\mathcal{H}(E_1, E_2, E_3)$ will give deviations from neutrality. Our indexes signal that ambiguity aversion and insensitivity are not uniform here. The basic design with the six combinations of parity and color can yield the average aversion/insensitivity over all events. □

Ambiguity is too rich a domain to expect that an agent will have one attitude for all events. There can be many kinds of (source) preferences and (lacks of) understanding of uncertainty beyond risk that go beyond whether probabilities are known or unknown (Tversky and Fox, 1995). Ambiguity attitudes depend on sources of uncertainty just like utility functions depend on commodities (Cappelli et al., 2020). Our indexes can examine such dependencies and emotions.

---

[10] Preferably, this involves a limited number of events, not only for tractability reasons, but also to stay away from very extreme a-neutral probabilities below 0.05 or above 0.95, as recommended elsewhere in this paper.

3.3. Uniform sources

We next investigate when different designs do give the same indexes.

Definition 15. The indexes *fit perfectly* if every measurement design $\mathcal{H}$ gives the same indexes. □

The following property characterizes perfect fit: m is *neo-additive* if there exist a probability measure P on S, $0 \leq \sigma \leq 1$, and $0 \leq \tau \leq 1 - \sigma$ such that

$$
\begin{aligned}
P(E) = 0 &\quad \Rightarrow \quad m(E) = 0; \\
0 < P(E) < 1 &\quad \Rightarrow \quad m(E) = \tau + \sigma P(E); \\
P(E) = 1 &\quad \Rightarrow \quad m(E) = 1.
\end{aligned}
\tag{10}
$$

W is *neo-additive* if the three implications in Eq. (10) hold with W instead of m, where furthermore $\sigma > 0$ and all $P(E_i) > 0$ (to satisfy monotonicity[11]). Under Assumptions 1, 2, and 8, and neo-additivity of m (Eq. (10)), substitution (Appendix B) gives:

$$
b = 1 - 2\tau - \sigma \quad \text{and} \quad a = 1 - \sigma.
\tag{11}
$$

As is common in axiomatizations, we assume complete information about preferences. That is, we consider all measurement designs and use m for all events.[12] As common in axiomatizations, we assume a continuum domain. We do so through the following conditions of Villegas (1964). We say that m is *fine*[13] if for each nonnull event A there exists an event $B \subset A$ such that $m(A) > m(B) > 0$. *Fineness* of P is defined similarly. *Event-continuity* holds if: (i) whenever a nested sequence $A_1 \supset A_2 \supset A_3 \supset \cdots$ converges to $\emptyset$ and $m(B) > 0$, there exists a J such that $m(A_j) < m(B)$ for all $j \geq J$, and (ii) whenever a nested sequence $B_1 \subset B_2 \subset B_3 \cdots$ converges to B and $m(B) > m(A)$, there exists a J such that $m(B_j) > m(A)$ for all $j \geq J$.

Theorem 16. [14] *Under Assumptions 1, 2, and 8, the following two statements are equivalent:*

(i) *m is neo-additive, and the corresponding probability measure P is fine ("atomless"[15]) and countably additive.*

(ii) *Our indexes perfectly fit, and m is fine and event-continuous.* □

By monotonicity, m in (i) is strictly increasing in P; i.e., $\sigma > 0$. The literature has documented several appealing properties of the neo-additive model (Eichberger et al., 2012, p. 238 penultimate paragraph). Theorem 16 provides another one and shows a new way to test the neo-additive model: by testing consistency of our indexes over different partitions.

---

[11] This also avoids some open mathematical problems in Chateauneuf et al. (2007), concerning nonnecessity of null event consistency in their Theorem 5.2 and inconsistency between null events in bets on and bets against events under their maximal pessimism.

[12] So far we, only used m on one fixed measurement design.

[13] We avoid the common term atomless because the term atom is used in reference to $\mathcal{H}$ in this paper. In the presence of the assumed event-continuity, our condition is equivalent to Savage's (1954) fineness. Generalizations that allow for atoms may be possible using Mackenzie's (2019) generalization of Villegas (1964).

[14] To avoid some Banach-Kuratowski-Ulam impossibility results, measure-theoretic structure can be added in this theorem, where the set of events is a σ-algebra.

[15] Here, atoms refer to S. In the rest of the paper, they refer to $\mathcal{H}$.

3.4. Applicability of our indexes

Theorem 16 shows that our indexes perfectly measure ambiguity attitudes if m is neo-additive. The neo-additive model performs well empirically and can capture the commonly observed four-fold pattern of ambiguity aversion (Trautmann and van de Kuilen, 2015).

Theorem 16 can be interpreted as the counterpart for ambiguity of a classical result in expected utility for risk: that the CRRA index perfectly captures (relative) risk aversion irrespective of the stimuli used if and only if utility is from the CRRA family (Harvey, 1990, Theorem 3). The CRRA index is tractable and performs well empirically, even though it usually does not fit perfectly due to nonconstant relative risk aversion and its difficulty in handling extreme outcomes. Despite these limitations, the CRRA index may still give the best ("average") summary of the data and capture risk attitudes well.

Likewise, we believe that our indexes capture ambiguity attitudes well even if they only approximate these. Our indexes do not fit all data well, e.g., as in Example 14, where the source of uncertainty is not "uniform" (formalized by Abdellaoui et al., 2011; our Eq. (22)). In such situations, however, no (pair of) indexes can fit all the data, and our indexes still give the best ("average") summary.

Our indexes may not work well either if very unlikely events are incorporated into the measurement design. Such events are known to involve many irregularities (Kahneman and Tversky, 1979; Ortoleva, 2012) and are better avoided in applications, e.g., by imposing boundary conditions (Tversky and Wakker, 1995; see Wakker's 2010 insensitivity region). Our indexes capture ambiguity attitudes well if practitioners take these limitations into account then.

Measuring risk attitudes using expected utility involves only one parameter: utility. However, measuring ambiguity attitudes using ambiguity models involves several parameters besides ambiguity attitudes, including utility, a-neutral probabilities, and risky-probability weighting. This complicates the task. Many preceding studies jointly estimated all those parameters to measure ambiguity attitudes. By contrast, we only use a few indifferences. Our indexes leave the other parameters free. They simply drop from the equations. Our indexes substantially simplify the measurement of ambiguity attitudes.

4. Extension to many outcomes

From now on, we drop Assumption 1 (only two outcomes) and consider general outcome sets X. Observation 17 extends the results obtained before without committing to a specific ambiguity model.

Observation 17. *All results of §2 and 3, including Theorems 11, 12, 16, and Corollary 13, remain valid if we drop Assumption 1 but fix two outcomes $\gamma \succ \theta$ for the matching probabilities m (Eq. (1)) and the indexes b, a. □*

5. Extension to outcome-dependent ambiguity models

In several models, ambiguity attitudes depend on the outcomes considered (e.g., Chew et al., 2008; Dobbs, 1991; Gul and Pesendorfer, 2014; Neilson, 2010; Siniscalchi, 2009; Skiadas, 2013). Then, the indexes in Observation 17 will depend on the outcomes γ, θ chosen, and can be used to investigate this dependence. For example, constant ambiguity aversion with respect to absolute utility increments (Grant and Polak, 2013), or proportional utility increments (Chateauneuf

and Faro, 2009), or wealth increments (Cerreia-Vioglio et al., 2019b) are inherited by our indexes and can be tested using them.

The most popular outcome-dependent ambiguity model is the smooth model (Klibanoff et al., 2005). We now analyze our indexes for this model. We assume that all functions are sufficiently smooth with all required derivatives existing and all O and o terms (defined later) uniform. Our analysis is similar to Izhakian and Brenner (2011) and Maccheroni et al. (2013), who provided local ambiguity premiums expressed in monetary units. Our premiums are expressed in probability units. The text below Eq. (5) explained that, under some assumptions, our index corresponds with a proportional monetary premium.

With fixed $\gamma \succ \theta$, the indifference $\gamma_E \theta \sim \gamma_{m(E)} \theta$ in Eq. (1) becomes under the smooth model of ambiguity, explaining notation below:

$$\int_{\Delta(S)} \varphi\big(Q(E)\big) d\mu = \varphi\big(m(E)\big). \tag{12}$$

The smooth model assumes expected utility for risk with utility function u, which we normalize at $u(\gamma) = 1$ and $u(\theta) = 0$. $\Delta(S)$ denotes the set of (first-order) probability measures over S, and μ is a second-order probability distribution over $\Delta(S)$ interpreted as perception of ambiguity. To evaluate $\gamma_E \theta$ (through the integral in Eq. (12)), we take the second-order μ-weighted expectation of $Q(E)$, the Q-expected utility, but transformed by a function φ. Concavity of φ captures ambiguity aversion, linearity captures ambiguity neutrality, and convexity captures ambiguity seeking. The right prospect in Eq. (1), $\gamma_{m(E)} \theta$, is evaluated by the right-hand side of Eq. (12) because the first-order probability of receiving γ is, certainly and unambiguously, $m(E)$. Let $p = P(E) = \int_{\Delta(S)} Q(E) d\mu$ denote the a-neutral probability of E. The variance of $Q(E)$ with respect to μ is $\sigma^2 = \int_{\Delta(S)} (Q(E) - P(E))^2 d\mu$. $A = -\frac{\varphi''}{\varphi'}$ is the Arrow-Pratt index of φ and captures ambiguity (Klibanoff et al., 2005, p. 1865). Further, σ^2 is the variance of the second-order uncertainty μ about ambiguity-neutral probabilities p, and $o(\sigma^2)$ expresses first-order approximation as σ^2 vanishes.

Observation 18. *In the smooth ambiguity model, the aversion index is*

$$b = \overline{\sigma^2 A(p)} + o\big(\sigma^2\big), \tag{13}$$

and the insensitivity index is

$$a = \frac{1}{2} \frac{Cov(\sigma^2 A(p), \nu)}{Var(\nu)} + o\big(\sigma^2\big). \quad \square \tag{14}$$

Thus, the ambiguity aversion index b is the (average of the) product of what is sometimes interpreted as ambiguity perception (σ^2) and a relative aversion index per perceived unit, $A(p)$. A similar decomposition occurs in Eq. (20) below, where it is discussed further. Remarkable is that Eq. (13) makes the average ambiguity aversion in the smooth model directly observable through b, even though its components p, A, and σ^2 are not.

The insensitivity index captures how the aversion premium ($\sigma^2 A(p)$ in Eq. (13)) increases with the event size ν, which indeed reflects sensitivity. This degree of ambiguity (perception), depending on the variance of the event probability, is similar to Izhakian's (2017) measure in his variation of the smooth model that uses Choquet expected utility instead of expected utility in the second stage.

The ambiguity attitude analyzed in the smooth model depends on the outcome interval $[\theta, \gamma]$ (e.g., Klibanoff et al., 2005, Proposition 4). It is independent of outcomes when $\varphi(x) = -e^{-\rho x}$ (Klibanoff et al., 2005, Proposition 2). Then:

$$b = \rho \overline{\sigma^2} + o(\sigma^2); \tag{15}$$

$$a = \frac{1}{2}\rho \frac{Cov(\sigma^2, \nu)}{Var(\nu)} + o(\sigma^2). \tag{16}$$

This case is of special interest because it concerns the intersection with the variational model (Maccheroni et al., 2006). This intersection is exactly the multiplier preference model of Hansen and Sargent (2001). Outcome independence is central in the next section.

6. Extension to outcome-independent ambiguity models

Many models assume that ambiguity attitudes are outcome independent.[16] Then our indexes are as well. The outcome-independent models are all special cases of *uniseparable utility*: there exists a worst outcome θ ($\forall \gamma \in X : \gamma \succcurlyeq \theta; \exists \gamma \succ \theta$) such that[17]

$$\gamma_E \theta \to W(E)U(\gamma) \quad \text{and} \quad \gamma_p \theta \to w(p)U(\gamma) \tag{17}$$

represents preferences for prospects with at most one outcome γ other than θ. For money, θ is usually set to 0. For health, it is typically set equal to death. Under prospect theory, θ is the reference outcome. U is a nonconstant *utility function*; we scale $U(\theta) = 0$. W is a nonadditive (*event*) *weighting function*, i.e., $W(\emptyset) = 0$, $W(S) = 1$, and W is *set-monotonic* ($A \supset B$ then $W(A) \geq W(B)$). Further, $w : [0, 1] \to [0, 1]$ is a (*probability*) *weighting function*, i.e., $w(0) = 0$, $w(1) = 1$, and w is strictly increasing. *Expected utility* implies (a) W is additive (i.e., W is a subjective probability measure) and (b) w is the identity. *Expected utility under risk* only implies (b). Under uniseparable utility, we can redefine m in the following outcome-independent manner:

Definition 19. $m(E) = p$ if $\gamma_E \theta \sim \gamma_p \theta$ for some $\gamma \succ \theta$. $\square$

By Eq. (17), Definition 19 is equivalent to $\gamma_E \theta \sim \gamma_p \theta$ for all $\gamma \succ \theta$ (and to $W(E) = w(p)$). It extends our preceding definition (Eq. (1)) to more than two outcomes. The choice of outcomes $\gamma \succ \theta$ is immaterial because they all give the same result and, consequently, Assumption 1 is no real restriction.

We could have increased outcome independence by imposing Savage's (1954) P4. This would allow θ to vary, as in biseparable utility, but would reduce generality by excluding some models.

For the sake of easy reference, we provide the following trivial reformulation of the results derived in preceding sections, adapted to general X and uniseparable utility.

---

[16] They include biseparable utility (Ghirardato and Marinacci, 2001), Choquet expected utility or rank-dependent utility, prospect theory for gains, maxmin EU, and the α-maxmin model (including Gul and Pesendorfer, 2015). Further included are separate-outcome weighting models ($\gamma_E \beta \to W(E)U(\gamma) + W(E^c)U(\beta)$; Einhorn and Hogarth, 1985), Chateauneuf and Faro's (2009) confidence representation with worst outcome θ, Izhakian's (2017) uncertain probability model, and Lehrer and Teper's (2015) event-separable representation.

[17] The assumption of a worst outcome is made or simplicity. Because we focus on gains, sign-dependence, as in prospect theory, plays no role.

Observation 20. *All results of §2 and §3 remain valid, including Theorems 11, 12, 16, and Corollary 13 if we replace Assumption 1 by uniseparable utility and use Definition 19 instead of Eq. (1). Ambiguity neutrality (Definition 3) then implies $W(.) = w(P(.))$ for a subjective probability measure $P (= m)$.* □

Observation 20 shows that our indexes and results can be applied to event-driven ambiguity models. It also shows that Definition 3 (ambiguity neutrality) agrees with common definitions. Thus, the results obtained in the first part of this paper for two fixed outcomes hold in great generality. Ambiguity neutrality comprises both probabilistic sophistication (Machina and Schmeidler, 1992) and indifference between subjective and objective probabilities (Dean and Ortoleva, 2017, Footnote 31).

7. Generalizing and unifying existing ambiguity indexes and orderings

This section applies our indexes to a number of outcome-independent ambiguity models and relates them to existing indexes and orderings. We assume that $\mathcal{H}$ is a belief hedge throughout.

7.1. Qualitative ambiguity orderings

To our best knowledge, ambiguity neutrality has always been defined as (a special case of) global probabilistic sophistication, sometimes as expected utility (surveyed by Gilboa and Marinacci, 2016). Then m is an additive probability (Definition 3 and Observation 20) and both our indexes are 0 (Theorem 11), which is compatible with the existing definitions. The sign of b then properly reflects ambiguity aversion/seeking.

It is common in the literature to define $\succcurlyeq^1$ as more ambiguity averse than $\succcurlyeq^2$ if $f \succcurlyeq^1 r \Rightarrow f \succcurlyeq^2 r$ where f is a general, possibly ambiguous act and r is an unambiguous act (risky, with known probabilities). See Dean and Ortoleva (2017, Definition 5), Gul and Pesendorfer (2014, Corollary 1), and Gul and Pesendorfer (2015, Propositions 3 and 4). Gul and Pesendorfer (2020) used a stronger restriction and defined the above condition as weakly more ambiguous. These conditions all imply that $\succcurlyeq^1$ has lower matching probabilities and, hence, a larger b index, which is again compatible with these definitions.

Some papers considered qualitative orderings of insensitivity or, relatedly, ambiguity perception. Multiple prior models have used set-inclusions of sets of priors for this purpose (Ghirardato et al., 2004, Proposition 6) which, for tractable subcases of multiple prior models, agrees with our insensitivity index (Eq. (20) below). Tversky and Wakker (1995) considered comparative subadditivity for general weighting functions W. If applied to matching probabilities, this agrees with Baillon and Bleichrodt's (2015) indexes (discussed in §7.2) and, therefore, with our index a. Similarly, Tversky and Wakker's (1995) source preference conditions agree without index b.

Some papers defined ambiguity indexes and orderings using premiums in monetary units rather than in our probability units (Brenner and Izhakian, 2018; l'Haridon et al., 2018; Maccheroni et al., 2013). These indexes depend on the utility function, are outcome-oriented, and are not directly related to our indexes. The remainder of this section shows that our indexes generalize many existing quantitative indexes.

7.2. Biseparable utility (including Choquet expected utility)

Many theories are special cases of uniseparable utility, including biseparable utility and Choquet expected utility with a nonadditive measure W. They often adopt an aversion index

$$1 - W(E) - W(E^c). \tag{18}$$

It was suggested by Schmeidler (1989, example on pp. 571-572 & p. 574) and proposed by Dow and Werlang (1992). Most theories assume expected utility for risk, so that $m = W$, and we get:

Observation 21. *Under Assumptions 2 and 8, expected utility for risk, and complementation-closedness of $\mathcal{H}$, our ambiguity aversion index b is the average of Eq. (18). In Schmeidler's (1989) model, ambiguity aversion[18] implies $b > 0$, ambiguity neutrality implies $b = 0$, and ambiguity seeking implies $b < 0$.* □

Theoretical studies often used Eq. (18) to define ambiguity aversion (Klibanoff et al., 2005, Definition 7). Empirically implementing it is complex because W needs to be known. Our aversion index shows how to make Eq. (18) observable without the need to measure W. Our indexes can do this without assuming expected utility for risk, which increases their descriptive validity.

Baillon and Bleichrodt (2015) considered a domain $\mathcal{H}(E_1, E_2, E_3)$ as in our Corollary 13, and measured five event-specific indexes. These indexes did not control for beliefs. Our indexes show how their indexes can be aggregated to provide that control, capturing both aversion and insensitivity.[19] Our indexes are also compatible with those of Chateauneuf et al. (2007).[20] Under Choquet expected utility with expected utility for risk, our Theorem 16 provides an alternative axiomatization of Chateauneuf et al. (2007) neo-additive model. Their model is in the intersection of Choquet expected utility and multiple prior models, to which we turn next.

7.3. Multiple priors

This subsection shows how common indexes of ambiguity in multiple prior theories can be measured directly for real-world applications without the need to measure utility or the set of priors or to make arbitrary assumptions about the set of priors. Let C denote a convex set of probability distributions over S. $P^*(E) = \sup_{P \in C} P(E)$ denotes *upper probabilities* and $P_*(E) = \inf_{P \in C} P(E)$ denotes *lower probabilities*. In the *α-maxmin* model (Ghirardato et al., 2004), preferences maximize, for $\gamma \succcurlyeq \beta$:

$$\gamma_E \beta \rightarrow W(E) U(\gamma) + (1 - W(E)) U(\beta)$$

---

[18] Schmeidler defined ambiguity aversion [neutrality; seeking] as quasiconvexity [linearity; quasiconcavity] of preference with respect to outcome (2nd stage probabilities) mixing, which implies positivity [nullness; negativity] of Eq. (18) for all E_i and, hence, of our b. He used the term uncertainty instead of ambiguity.

[19] Using their notation: $b = \overline{BC}$ and $a = \overline{(LA + UA)}/3$.

[20] The authors' interpretations strongly suggest expected utility for risk, and we assume it. (Without this assumption, their indexes do not solely capture ambiguity attitudes but also risk attitudes.) Then $m = W$ is neo-additive and Eq. (11) gives our indexes. Chateauneuf et al. (2007, p. 544 top) interpret a (we use our notation) as lack of confidence (or distrust) in the a-neutral probability P, and $\frac{b}{2a} + \frac{1}{2}$ as an index of pessimism. Ignoring the irrelevant term $\frac{1}{2}$, their pessimism index is our aversion per unit of distrust in P, which is a relative analog of our absolute index. We compare such relative and absolute versions after Eq. (20) below.

with $W(E) = \alpha P_*(E) + (1 - \alpha)P^*(E)$ $(0 \leq \alpha \leq 1)$. Expected utility is assumed for risk. Maxmin expected utility is the special case of $\alpha = 1$ (Alon and Schmeidler, 2014). Under complementation-closedness we obtain (proved in Appendix B; v-hedging is not needed here):

$$b = (2\alpha - 1)\overline{\left(P^* - P_*\right)}. \tag{19}$$

Here $\overline{\left(P^* - P_*\right)}$ is the average discrepancy between upper and lower probabilities of events, which is sometimes interpreted as ambiguity perception—or as the Dempster-Shafer plausibility-belief gap (Gul and Pesendorfer, 2014, Corollary 2; Gul and Pesendorfer, 2020, p. 7; Walley, 1991, p. 222). Further, $2\alpha - 1$ (or, equivalently, α itself) is commonly taken as an index of ambiguity aversion. It is 0 under ambiguity neutrality.

A tractable subclass of α-maxmin is the ε-α-*maxmin model*. Let $C = \{(1 - \varepsilon)Q + \varepsilon T\}$, with a fixed baseline probability Q, a fixed $\varepsilon \in [0, 1]$, and the variable T any probability measure (Dimmock et al., 2015; axiomatized by Chateauneuf et al., 2007). It is a subclass of the ε-contamination model (Ellsberg, 1961, pp. 663-669) that has been used in many fields (e.g., Aryal and Stauber, 2014; Epstein and Schneider, 2010; Hodges and Lehmann, 1952). Several authors proposed ε $(= \overline{P^* - P_*})$, which captures the size of the set of priors, as an index of ambiguity perception (e.g., Alon and Gayer, 2016; Chateauneuf et al., 2007, p. 543; Ghirardato et al., 2004 Proposition 6; Walley, 1991). Dimmock et al. (2015) showed:

$$a = \varepsilon \quad \text{and} \quad b = (2\alpha - 1)\varepsilon. \tag{20}$$

Thus, our insensitivity index directly agrees with ambiguity perception. Index α (or $2\alpha - 1$) captures ambiguity aversion in a relative sense, as aversion per perceived unit of ambiguity. Our index b is the product of what is often interpreted as ambiguity perception and aversion per unit of perception, and is a measure of absolute ambiguity aversion. The pairs (a, b) and (ε, α) are informationally equivalent, and which pair is most convenient depends on the context. Index b is most useful for determining ambiguity premiums.[21]

Hey et al. (2010) used the following special case of α-maxmin. They considered three atoms E_1, E_2, E_3, and C contained all P with $P(E_1) \geq \varepsilon_1$, $P(E_2) \geq \varepsilon_2$, $P(E_3) \geq \varepsilon_3$, where the ε_j are nonnegative and sum to less than 1. Then, with similar interpretations as before[22]:

$$a = 1 - \varepsilon_1 - \varepsilon_2 - \varepsilon_3 \quad \text{and} \quad b = (2\alpha - 1)a. \tag{21}$$

Dimmock et al. (2015) showed how to make ambiguity aversion and ambiguity perception directly observable, without the need to measure utility U or the set of priors C, for Ellsberg urn events. They used Ellsberg urn events and assumed expected utility for risk. Our indexes allow us to extend their analysis to real-world events and drop the expected utility assumption for risk.

7.4. The source method

Abdellaoui et al.'s (2011) source method is the specification of Choquet expected utility with

$$W(E) = w_{So}\big(P(E)\big). \tag{22}$$

Here w_{So} is strictly increasing with $w_S(0) = 0$ and $w_S(1) = 1$, and P designates a-neutral probabilities. The subscript *So* expresses dependence on the source of uncertainty, and this dependency

---

[21] Schmeidler (1989 p. 574) used the term uncertainty premium for index b.

[22] This follows from Eq. (20) by defining $Q(E_j) = \frac{\varepsilon_j}{\varepsilon_1 + \varepsilon_2 + \varepsilon_3}$ and $\varepsilon = 1 - \varepsilon_1 - \varepsilon_2 - \varepsilon_3$.

captures ambiguity as, for instance, in Ellsberg's paradoxes. Abdellaoui et al. (2011) call a source *So uniform* if Eq. (22) is satisfied. We focus here on one uniform source *So* of ambiguity—besides risk with known probabilities.

Abdellaoui et al. (2011) and Dimmock et al. (2016), AD henceforth, fitted neo-additive forms of w_S and $m(E)$, respectively, on the open interval $(0, 1)$. They then derived their indexes from this. We here do so for the function $m(E)$ (Eq. (10)), where $\sigma \geq 0$ and τ are chosen to minimize distance. To do so, they needed to know the a-neutral probabilities $P(E)$. Dimmock et al. (2016) solved this by making the common assumption of symmetry of colors in Ellsberg urns, whereas Abdellaoui et al. (2011) measured the probabilities separately. With τ and σ the best-fitting neo-additive parameters, AD defined (as in Eq. (11))

$$b' := 1 - 2\tau - \sigma, a' := 1 - \sigma. \tag{23}$$

By the properties of linear regression estimators we get: $a' = 1 - \frac{Cov(m, P)}{Var(P)}$ and by Eq. (9) and the results in Appendix A, our index a agrees well with a'. Our index b always agrees with b'. Again, we contribute to the literature by showing how ambiguity attitudes can be measured without knowing or making assumptions about P.

8. Discussion

Although our indexes can be used beyond the widely studied Ellsberg urns, it remains interesting to apply them to these. Example 14 is a variation of the well-known three-color Ellsberg urn, where different sources of uncertainty come together. Studying such situations is an interesting topic for future research, both empirically and theoretically (Eliaz and Ortoleva, 2016). Cappelli et al. (2020) give some theoretical suggestions.

Recently, much attention has been paid to "hedging" possibilities as a confound in ambiguity experiments (Agranov and Ortoleva, 2017; Bade, 2015; Baillon et al., 2020; Cerreia-Vioglio et al., 2019a; Dean and Ortoleva, 2017; Georgalos, 2019; Oechssler et al., 2019). Subjects may use ambiguities in some choices in the experiment to "hedge" against ambiguities in other choices in the experiment. This concept of hedging is unrelated to our concept. Given the level of sophistication (which usually includes knowledge of the experimental design) required for such hedging, it is more likely that subjects treat each choice in isolation (Binmore et al., 2012, p. 229; Georgalos, 2019, p. 57; Oechssler et al., 2019; Starmer and Sugden, 1991), so that the problem does not arise. However, even if only a few subjects hedge, then this confounds experiments on ambiguity and, in fact, any preference experiment. It is, therefore, desirable to maximally enhance isolated perception in choice experiments. Baillon et al. (2020) and Johnson et al. (2021) provided methods for doing so.

We, finally, use our running example to illustrate some advantages of our approach.

Example 22 (*Example 4 Continued*). Eichberger and Kelsey (2011), EK henceforth, also considered marginal cost $c = 0.1$ besides $c = 0.9$. The observed changes in choice percentages (shifting towards high effort) were intuitive, but hard to explain by classical game theory. EK showed that they can be explained by ambiguity theories. For empirically plausible ambiguity attitudes, they referred to another paper, Kilka and Weber (2001), who used different sources of uncertainty (and subjects) and inferred subjective beliefs from introspective judgments. Hence, it was not revealed-preference based. In terms of our indexes, they found as plausible ranges: $-0.15 \leq b \leq 0.12$ and $0.41 \leq a \leq 0.61$ (EK p. 319). These include the values ($b = 0.08$, $a = 0.43$) that we obtained in our imaginary example.

Using belief hedges, we can measure ambiguity attitudes with the following advantages: (1) they are the attitudes of the players themselves; (2) they refer directly to the relevant uncertainty (the effort level of the other player); (3) the a-neutral probabilities of the players need not be known to us—players do not know the percentages in Table 1; (4) we use only revealed preferences. An additional advantage of the basic design used here, which involved all relevant uncertainties (s_j), is that we can derive estimates of the underlying a-neutral probabilities. For instance, for our imaginary player, we get $p_1 = 0.54$, $p_2 = \cdots = p_5 = 0.07$, $p_6 = 0.19$. □

Given our indexes and belief hedges, it is trivial to see that Baillon et al. (2018) is a special case. Our contribution concerns the reversed direction: given Baillon et al.'s results (our Corollary 13), we develop the general indexes and the concept of belief hedges. Finding Eq. (8) as the proper general concept of insensitivity, was the most challenging step. The validity of the general indexes was subsequently confirmed by theoretical justifications: preference axiomatizations and the common generalization of virtually all existing indexes, including Baillon et al.'s.[23] Another challenge was to find the concept of belief hedges needed for the required flexibility and tractability of our aversion and insensitivity indexes in applications (Examples 4, 14, and 22). In many real-world situations, control over the data is limited, and the flexibility of general belief hedges is desirable. Other practical advantages over the special case of Baillon et al. (2018) were discussed at the beginning of §3.

9. Conclusion

For a long time, ambiguity measurements were confined to artificial events such as opaque urns because it was unknown how to control for unknown beliefs. We introduced belief hedges to measure ambiguity attitudes when subjective beliefs are unknown. Belief hedges extend the hedging concept from finance, where it protects against unknown outcomes, to ambiguity where it protects against unknown beliefs. We show that belief hedges are necessary and sufficient for measuring ambiguity attitudes when beliefs are unknown. This allows us to measure ambiguity attitudes in real-world applications.

Using belief hedges and some econometric concepts, we introduce two new indexes of ambiguity. They can easily be applied in practice (Examples 4, 14, and 22), and they generalize most other indexes proposed in the literature, including those of Baillon et al. (2018). They thus unify existing indexes, including ambiguity orderings. Our indexes are valid under virtually all existing ambiguity theories and do not require expected utility for risk or multi-stage stimuli, which makes it easy to use them and which is needed for empirical validity. They can also accommodate the empirically observed ambiguity seeking for unlikely events. They use no theoretical constructs and can be directly revealed from preferences. In this sense, they operationalize earlier indexes. Our indexes need no measurements and data fittings of risk attitudes (utility/probability weighting) or a-neutral probabilities. They make ambiguity theories widely applicable.

---

[23] Baillon et al. (2018) did not provide theoretical justifications.

Appendix A. Goodness of fit of Eq. (9)

Throughout this paper, for $F : \mathcal{H} \rightarrow \mathfrak{R}$, we write[24]: $\overline{F} = \frac{\sum_{E \in \mathcal{H}} F(E)}{|\mathcal{H}|}$; $Var(F) = \frac{\sum_{E \in \mathcal{H}} (F(E) - \overline{F})^2}{|\mathcal{H}|}$; $Cov(F, G) = \frac{\sum_{E \in \mathcal{H}} (F(E) - \overline{F}) \times (G(E) - \overline{G})}{|\mathcal{H}|}$.

The following lemma considers variations within our constructed domain $\mathcal{H}$.

Lemma 23. *Assume Assumption 8 and l-hedging. Equivalent are:*

> (i) *v-hedging;* (ii) $E(1_s \times v)$ *is the same for each s;* (iii) $Cov(1_s, v)$ (24)
>
> *is the same for each s.*

Now, also assume v-hedging. With 1_{E_i} the probability measure on the atoms assigning probability 1 to E_i, we have, for all s, i, P:

$$\frac{Cov(1_s, v)}{Var(v)} = \frac{Cov(1_{E_i}, v)}{Var(v)} = \frac{Cov(P, v)}{Var(v)} = \frac{Cov(v, v)}{Var(v)} = 1. \qquad (25)$$

Proof. For Eq. (24), (ii) is a rewriting of (i), and (iii) is equivalent because $Cov(1_s, v) = (E(1_s \times v) - E(1_s) \times E(v)) = E(1_s \times v) - \frac{1}{4}$.

For Eq. (25), the first fraction is the same for all s by Eq. (24)(iii). The first equality now follows because $1_s = 1_{E_i}$ on $\mathcal{H}$ for each $s \in E_i$. The second equality follows because every probability measure P on $\mathcal{H}$ is a convex combination of measures $1_{E_i}(.)$, and sensitivity and covariance are compatible with convex combinations.[25] The third equality follows because v is a special case of a probability measure, and the last equality is by definition. $\square$

We next turn to extraneous randomness in the dependency of m on P and v. The above equality $\frac{Cov(P, v)}{Var(v)} = 1$ means that, on average, a change of one unit of v generates one unit change of P. Hence, by Stock and Watson (2015, §12.1 and Eq. (12.7)), Eq. (9) provides the best first-order approximation under common econometric assumptions together with the following critical assumption: m depends on v only through P with, further, random noise. To illustrate this result, and explain when the approximation works well, we give an independent derivation of the following result.

Proposition 24. *Under Assumptions 2 and 8 and belief hedging, Eq. (9) holds with exact equality if any of the following three conditions holds:*

> (i) $P = v$;
> (ii) *m is neo-additive;*
> (iii) *P best fits m.*[26]

---

[24] We use population statistics. If one interprets $\mathcal{H}$ as a sample, small relative to $|S|$, then one may prefer sample statistics, with denominators $|\mathcal{H}| - 1$ instead of $|\mathcal{H}|$. However, those always give the same indexes and results throughout our paper because the denominator cancels from all equations.

[25] That is, the sensitivity (or covariance) of a convex combination of functions with respect to some other variable (v in our case) is the convex combination of their sensitivities (or covariances).

[26] We take the neo-additive model that minimizes quadratic distance, as common in regressions.

Proof. (i) is trivial, and (ii) follows from Eq. (11) (irrespective of what P is). We consider (iii), where m is related to P through the neo-additive decision model ("regular regression"). The distance to be minimized is

$$\sum_{E \in \mathcal{H}} \big(m(E) - \tau - \sigma P(E) \big)^2. \tag{26}$$

The first order condition of Eq. (26) with respect to τ, divided by -2, gives $\sum_{E \in \mathcal{H}} (m(E) - \tau - \sigma P(E)) = 0$. Thus, using Eq. (3),

$$\tau = \overline{m} - \sigma/2. \tag{27}$$

In words, the best-fitting line passes through the center of gravity of the data points, being $(\frac{1}{2}, \overline{m})$. We define the additive measure $Q(E) := \sigma P(E)$ and $q_i := Q(E_i) = \sigma P(E_i)$ and find the optimally fitting q_i. We optimize over all $q_i \in \mathbb{R}$, later verifying that they are all positive (and $\sigma > 0$). By Eq. (27), the distance to be minimized becomes

$$\sum_{E \in \mathcal{H}} \big((m(E) - \overline{m}) - (Q(E) - \sigma/2) \big)^2. \tag{28}$$

The first-order condition with respect to q_i is

$$\sum_{E \supset E_i} \big((m(E) - \overline{m}) - (Q(E) - \sigma/2) \big) = 0. \tag{29}$$

Summing over i:

$$\sum_i \sum_{E \supset E_i} \big((m(E) - \overline{m}) - (Q(E) - \sigma/2) \big) = 0. \tag{30}$$

$$\sum_{E \in \mathcal{H}} \left((m(E) - \overline{m}) - \sigma \left(P(E) - \frac{1}{2} \right) \right) v(E) = 0. \tag{31}$$

$$\sum_{E \in \mathcal{H}} \left((m(E) - \overline{m}) - \sigma \left(P(E) - \frac{1}{2} \right) \right) \left(v(E) - \frac{1}{2} \right) = 0. \tag{32}$$

$$\sigma = \frac{\sum_{E \in \mathcal{H}} (m(E) - \overline{m})(v(E) - \frac{1}{2})}{\sum_{E \in \mathcal{H}} (P(E) - \frac{1}{2})(v(E) - \frac{1}{2})} = \frac{|\mathcal{H}|}{|\mathcal{H}|} \frac{Cov(m, v)}{Cov(P, v)} = \text{(by Eq. (25))} \frac{Cov(m, v)}{Var(v)}. \tag{33}$$

By Eq. (25), the above denominators are positive. By monotonicity, the above numerators are positive; $\sigma > 0$; $q_i = \sigma p_i > 0$ for all i. Because, with P given, optimal fitting entails a regular regression of m w.r.t. P, it is well-known that $\sigma = \frac{Cov(m, P)}{Var(P)}$. Combining this with Eq. (33) implies exact equality in Eq. (9).   □

Eq. (9) gives a good approximation if any of the three cases in Proposition 24 holds approximately. Poor approximation can result if all these assumptions are strongly violated, but such cases are not empirically plausible. Poor approximation can, of course, also result if our basic assumptions, such as monotonicity, are violated. Eqs. (9) and (8) work well for all practical purposes.

Appendix B. Proofs except of Theorem 16 and for §5

Proof of Theorem 11. Under ambiguity neutrality, m is a probability measure on $\mathcal{H}$ and its atoms. By Eq. (3), $\overline{m} = 0.5$ and $b = 0$. By Eq. (25), $\frac{Cov(m,v)}{Var(v)} = 1$ and $a = 0$. Conversely, assume $b = 0$ for all probability measures $P = m$. Then $\overline{m} = 0.5$ for all $m = 1_s$, which is l-hedging. Similarly, if $a = 0$ for all probability measures $P = m$ then it is so for all $m = 1_s$, implying $\frac{Cov(1_s,v)}{Var(v)} = 1$ for all s which, by Eq. (24), implies v-hedging.

We, finally, turn to the supremum values of the indexes. b tends to its supremum 1 as $\overline{m}$ tends to its minimum 0. a tends to its supremum 1 as $Cov(m, v)$ tends to its infimum 0 (by monotonicity, it cannot be negative), which occurs when m tends to a constant function. $\square$

Proof of Theorem 12. We take our data set m as a vector in $\mathbb{R}^{|\mathcal{H}|}$. Index b is a normalization of the inner product of m with the *aversion vector* $(1, \ldots, 1)$. Index a is a normalization of the inner product of m with the *insensitivity vector* $(v(E) - \frac{1}{2})_{E \in |\mathcal{H}|}$.[27] The aversion and insensitivity vectors are orthogonal because their inner product is $\sum (v(E) - \frac{1}{2}) = 0$. Having inner product 0 is the formal definition of orthogonality and is equivalent to the geometric concept of rectangularity. $\square$

Proof of Corollary 13. The case of index b is clear. As for index a,

$$
|\mathcal{H}|Cov(m, v) = \sum_{i=1}^{3} \left(m(E_i) - \overline{m} \right) \left(\frac{1}{3} - \frac{1}{2} \right) + \sum_{i=1}^{3} \left(m(E_i^c) - \overline{m} \right) \left(\frac{2}{3} - \frac{1}{2} \right)
$$

$$
= 3(\overline{m_s} - \overline{m}) \left(-\frac{1}{6} \right) + 3(\overline{m_c} - \overline{m}) \left(\frac{1}{6} \right) = \frac{\overline{m_c} - \overline{m_s}}{2};
$$

$$
|\mathcal{H}|Var(v) = \sum_{i=1}^{3} \left(\frac{1}{3} - \frac{1}{2} \right)^2 + \sum_{i=1}^{3} \left(\frac{2}{3} - \frac{1}{2} \right)^2 = \frac{1}{6}.
$$

$$
\frac{Cov(m, v)}{Var(v)} = 3(\overline{m_c} - \overline{m_s}).   \square
$$

Proof of Eq. (11). Because $\emptyset$ and S are not in $\mathcal{H}$, and all atoms are nonnull, $0 < P(E) < 1$ for all $E \in \mathcal{H}$. Hence, $\overline{m} = \tau + \sigma \overline{P} = $ (by Eq. (3))$\tau + \sigma/2$ and the result for b follows. As regards a, because m is an affine function of P with slope σ on $\mathcal{H}$, $\frac{Cov(m,v)}{var(v)} = \sigma \frac{Cov(P,v)}{var(v)} = $ (by Eq. (25)) σ. $\square$

Proof of Eq. (19). We also assume complementation-closedness here (v-hedging is not needed here). $m(E^c) = \alpha P_*(E^c) + (1 - \alpha)P^*(E^c) = \alpha(1 - P^*(E)) + (1 - \alpha)(1 - P_*(E))$. Further, $m(E) + m(E^c) = \alpha P_*(E) + (1 - \alpha)P^*(E) + \alpha(1 - P^*(E)) + (1 - \alpha)(1 - P_*(E)) = 1 - (2\alpha - 1)(P^*(E) - P_*(E))$. Finally, $= 1 - \overline{2m(E)} = 1 - \overline{m(E) - m(E^c)} = (2\alpha - 1) \times \overline{(P^* - P_*)}$. $\square$

---

[27] $|\mathcal{H}|Cov(m, v) = \sum (m(E) - \overline{m})(v(E) - \frac{1}{2}) = \sum m(E)(v(E) - \frac{1}{2})$.

Appendix C. Proof of Theorem 16

That (i) implies (ii) in Theorem 16 follows because Eq. (11) holds for every $\mathcal{H}$. From now on, we assume (ii) and derive (i). To prepare, we first prove that, if our indexes fit perfectly, then we must have probabilistic sophistication within our source S. That is, we must have uniformity in the terminology of Abdellaoui et al. (2011), ruling out Example 14.

Observation 25. _Under Assumptions 2 and 8, if our indexes are the same for every $\mathcal{H}\{E_1, E_2, E_3\}$, and fineness and event-continuity hold, then $m(.) = w_a(P(.))$ for a strictly increasing w_a and a fine (atomless) countable additive probability measure P._

Proof. The proof uses Lemmas 26-30.

Lemma 26. _We cannot have $A_1 \succ B_1$, $A_2 \succcurlyeq B_2$, $A_3 \succcurlyeq B_3$ for two threefold partitions $\{A_1, A_2, A_3\}$ and $\{B_1, B_2, B_3\}$ of S containing nonnull events._

Proof. Consider $\mathcal{H}\{A_1, A_2, A_3\}$ and $\mathcal{H}\{B_1, B_2, B_3\}$. They have the same aversion index b and, hence, the same average $\overline{m}$. Because $\overline{m_s}$ of the former exceeds $\overline{m_s}$ of the latter, for $\overline{m_c}$ it must be opposite. But then (Corollary 13) a is smaller for the former than for the latter, contradicting perfect fit.   QED

We next derive implications of event continuity, similar to Villegas (1964, p. 1790) but we do not have what he called monotonicity ($\approx$ additivity)—this is also the reason that we need two event continuity conditions, whereas for Villegas one is equivalent to the other.

Lemma 27. _If $D \succ B \succ \emptyset$, then there exist $C \subset D$, $A \subset D$ with $D \succ C \succ B \succ A \succ \emptyset$._

Proof. There exists $H \subset D$ such that $D \succ H \succ \emptyset$. $D - H$ is nonnull and, by monotonicity, $\succ \emptyset$. We have partitioned D into two nonnull events that we now denote D_1, S_1, where we assume $D_1 \succcurlyeq S_1$. We can similarly partition the smaller of these two, S_1, into two nonnull events $D_2 \succcurlyeq S_2$, and inductively continue to obtain an infinite decreasing (in terms of $\succcurlyeq$) sequence of disjoint nonnull subevents $D_j \subset D$.

Assume, for contradiction, that $D_j \succcurlyeq B$ for all j, which can be interpreted as a violation of Archimedeanity. Whereas $\bigcup_{i=j}^{\infty} D_i$ decreases to the empty set for $j \to \infty$, every union is $\succ B \succ \emptyset$, violating event continuity. Hence, an $A = D_j$ as required exists. This also implies that $S_\infty := D - \bigcup_{i=1}^{\infty} D_i$ is null. Otherwise, with S_∞ in the role of B, $D_j \prec S_\infty$ should occur for some j as we just showed, contradicting $D_j \succcurlyeq S_j$. We can, therefore, replace D_1 by $D_1 \cup S_\infty$ and every S_j by $S_j - S_\infty$, without affecting preference. That is, $\bigcup_{i=1}^{\infty} D_i = D$. By event continuity, $C := \bigcup_{i=1}^{J} D_j \succ B$ for J large enough.   QED

Lemma 28. _We cannot have $A_1 \succ B_1$, $A_2 \succcurlyeq B_2$ for two twofold partitions $\{A_1, A_2\}$ and $\{B_1, B_2\}$ of S._

Proof. Assume, for contradiction, events as in the lemma. By Lemma 27, there exists $A_1' \subset A_1$ such that $A_1 \succ A_1' \succ B_1$. We define $A_1'' = A_1 - A_1' \succ \emptyset$ (by monotonicity). Again by Lemma 27, there exists $B_1'' \subset B_1$ with $\emptyset \prec B_1'' \prec A_1''$. We define $B_1' = B_1 - B_1''$. We have two partitions $\{A_1', A_1'', A_2\}$ and $\{B_1', B_1'', B_2\}$ that violate Lemma 26.   QED

Lemma 29. *If $A \cap C = B \cap C = \emptyset$, then $A \succcurlyeq B \Leftrightarrow A \cup C \succcurlyeq B \cup C$.*

Proof. Assume $A \succcurlyeq B$. Consider partitions $\{A, C, S - A - C\}$ and $\{B, C, S - B - C\}$. By Lemma 26, $S - A - C \preccurlyeq S - B - C$. By Lemma 28, $A \cup C \succcurlyeq B \cup C$. The same reasoning holds with strict preferences. QED

Villegas used the following implication.

Lemma 30. *Assume $A_1 \cap A_2 = \emptyset = B_1 \cap B_2$. Then $A_1 \succcurlyeq B_1$ & $A_2 \succcurlyeq B_2 \Rightarrow A_1 \cup A_2 \succcurlyeq B_1 \cup B_2$, with strict preference if at least one of the two premises is strict.*

Proof. By Lemma 29, and Villegas (1964, p. 1789 4th para). QED

Observation 25 now follows from Villegas (1964, Theorem 4.3). □

Observation 31. *m is neo-additive.*

Proof. By perfect fit, each belief hedge $\mathcal{H}$ imposes two equalities on $m(.) = w_a(P(.))$, one for each index. We know that there exists at least one w_a satisfying all those equalities, being the neo-additive function corresponding with the values b, a found (Eq. (11)). It, hence, suffices to show that $w_a(p)$ is uniquely determined for each p. Consider $\mathcal{H}\{E_1, E_2, E_3\}$ with $P(E_j) = \frac{1}{3}$ for each j. By fineness and countable additivity, such E_js exist. Here, b determines the average of $\overline{m_s} = w_a(\frac{1}{3})$ and $\overline{m_c} = w_a(\frac{2}{3})$ and a determines their difference. This uniquely determines $w_a(\frac{1}{3})$ and $w_a(\frac{2}{3})$ as the neo-additive values.

Next assume, for induction w.r.t. $k \geq 0$, that w_a takes the neo-additive values at all $p = \frac{i}{3 \times 2^k}$. Consider $\frac{j}{3 \times 2^{k+1}}$ ($< 1/2$) for an odd $j < 3 \times 2^k$, and a threefold partition $\{E_1, E_2, E_3\}$ with $P(E_1) = P(E_2) = \frac{j}{3 \times 2^{k+1}}$, so that $P(E_3) = \frac{3 \times 2^k - j}{3 \times 2^k}$. For $\mathcal{H}\{E_1, E_2, E_3\}$'s m values, there are only two unknowns: $w_a(\frac{j}{3 \times 2^{k+1}})$ (for E_1 and E_2) and $w_a(1 - \frac{j}{3 \times 2^{k+1}})$ (for $E_1 \cup E_3$ and $E_2 \cup E_3$). Again, Corollary 13 uniquely determines the average and the difference of the two unknowns, so that they are both uniquely determined and must be the neo-additive values. This way, w_a takes the neo-additive values at all $p = \frac{j}{3 \times 2^{k+1}}$, both below and above $\frac{1}{2}$. By induction, it does so for all k. These values lie dense in $(0, 1)$, so that the nondecreasing (by monotonicity it is even strictly increasing) function w_a is the neo-additive function everywhere. □

The following observation follows from the above proof because we only used the designs mentioned.

Observation 32. *Perfect fit in Statement (ii) in Theorem 16 can be restricted to designs $\mathcal{H}\{E_1, E_2, E_3\}$.* □

Appendix D. Proofs for the Smooth Model (§5)

Using the notation of Section 5, we first derive the analog of Pratt's (1964) Eq. (5) regarding his risk premium, which is in monetary units. Our ambiguity premium instead is in probability units. The following lemma illustrates once more that our treatment of uncertainty about probabilities is analogous to traditional treatments of uncertainty about outcomes.

Lemma 33. *For some given event E:*

$$p - m(E) = \frac{1}{2}\sigma^2 A(p) + o(\sigma^2).$$
(34)

Proof of Lemma 33. Pratt (1964, Eqs. (4)-(6)) studied local risk premiums by letting lotteries converge to a riskless lottery/outcome x, with expectation kept fixed and variance tending to 0. We similarly study local ambiguity premiums by letting acts converge to an unambiguous act/lottery $\gamma_p\theta$, with the ambiguity-neutral part kept fixed and ambiguity σ^2 tending to 0, as follows.

In our mathematical derivation we will use a mathematical extension of m, i.e. m as it *would be* in the smooth model for events derived from each $F \in \mathcal{H}$ as in Eq. (35) below (required for all $\alpha > 0$ sufficiently close to 0, where "sufficiently close" may depend on F). Such events need not be present in the actual design $\mathcal{H}$.

We follow Klibanoff et al. (2005) and assume a compound state space $S = S' \times (0, 1]$, providing an Anscombe-Aumann mixture structure. Here S' captures the uncertainty of interest and $[0, 1]$ is only auxiliary. For example, F' is the event of the AEX index going up by more than 0.2%, and $F = F' \times [0, 1]$ is the event of that happening and the result of our randomizing machine just being anything. F and F' can be identified for many purposes. In what follows, we keep some F and the corresponding F' fixed, with fixed a-neutral probability p (μ-averaged $Q(F)$) and fixed μ-variance of $Q(F)$, denoted τ^2. We consider mixtures $\alpha\gamma_F\theta + (1-\alpha)\gamma_p\theta$ comprising an α ambiguous and a $1 - \alpha$ unambiguous part, with $\alpha \downarrow 0$. This mixture can be obtained by receiving γ under the disjoint union of an ambiguous and unambiguous event:

$$(F' \times (1-\alpha, 1]) \cup (S' \times (0, (1-\alpha)p]);$$
(35)

and θ otherwise. The events in Eq. (35) play the role of events E in Eq. (34). The limit of E tending to an ambiguity neutral event in the main text is achieved by letting α tend to 0 in Eq. (35). The corresponding ambiguity-neutral probability is $\alpha p + (1-\alpha)p = p$ for all α.

The matching probability m_α is defined by the indifference

$$\gamma_{(F' \times (1-\alpha, 1]) \cup (S' \times (0, (1-\alpha)p])}\theta \sim \gamma_{m_\alpha}\theta.$$

Writing $q = Q(F)$,

$$\int_{\Delta(S)} \varphi(\alpha q + (1-\alpha)p)d\mu = \varphi(m_\alpha).$$
(36)

Substituting Taylor series of φ for $\alpha \downarrow 0$ in the right-hand side:

$$\varphi(m_\alpha) = \varphi(p) + (m_\alpha - p)\varphi'(p) + O((m_\alpha - p)^2)$$
(37)

and for the integrand of the left-hand side:

$$\varphi(\alpha q + (1-\alpha)p) = \varphi(p) + \alpha(q-p)\varphi'(p) + \frac{1}{2}\alpha^2(q-p)^2\varphi''(p) + o(\alpha^2).$$

Hence the left-hand side of Eq. (36) is:

$$\varphi(p) + \alpha\varphi'(p)\int_{\Delta(S)} (q-p)d\mu + \frac{1}{2}\alpha^2\varphi''(p)\int_{\Delta(S)} (q-p)^2 d\mu + o(\alpha^2)$$

$$= \varphi(p) + \frac{1}{2}\alpha^2\varphi''(p)\tau^2 + o(\alpha^2)$$
(38)

(the term with φ' drops). Because of Eq. (36) we can equate Eqs. (37) and (38):

$$(m_\alpha - p)\varphi'(p) + O\big((m_\alpha - p)^2\big) = \frac{1}{2}\alpha^2\varphi''(p)\tau^2 + o\big(\alpha^2\big).$$

Dividing by $\varphi'(p)$, which does not affect O or o:

$$(m_\alpha - p)\big(1 + O(m_\alpha - p)\big) = \frac{1}{2}\alpha^2\frac{\varphi''(p)}{\varphi'(p)}\tau^2 + o\big(\alpha^2\big).$$

$$m_\alpha - p = \frac{-\frac{1}{2}\alpha^2 A(p)\tau^2 + o(\alpha^2)}{(1 + O(m_\alpha - p))}$$

$$= -\frac{1}{2}\alpha^2 A(p)\tau^2 + \frac{O(m_\alpha - p)\frac{1}{2}\alpha^2 A(p)\tau^2 + o(\alpha^2)}{(1 + O(m_\alpha - p))}$$

$$= -\frac{1}{2}A(p)\alpha^2\tau^2 + o\big(\alpha^2\big) = -\frac{1}{2}A(p)\alpha^2\tau^2 + o\big(\alpha^2\tau^2\big).$$

$\alpha^2\tau^2$ here is the variance of the event in Eq. (35) i.e., it is denoted σ^2 in Eq. (34) which now follows. □

Proof of Observation 18. Index b is Eq. (34) averaged over $\mathcal{H}$ and multiplied by 2 for normalization (see §2.4). As for the insensitivity index, by Eq. (34), $a = 1 - \frac{Cov(p,v)}{Var(v)} + \frac{Cov(\frac{1}{2}\sigma^2 A(p) + o(\sigma^2), v)}{Var(v)}$. Because $Cov(p, v) = Var(v)$ (Eq. (25)), we obtain Eq. (14). □

References

Abdellaoui, Mohammed, Baillon, Aurélien, Placido, Laetitia, Wakker, Peter P., 2011. The rich domain of uncertainty: source functions and their experimental implementation. Am. Econ. Rev. 101, 695–723.

Agranov, Marina, Ortoleva, Pietro, 2017. Stochastic choice and preferences for randomization. J. Polit. Econ. 125, 40–68.

Alon, Shiri, Gayer, Gabrielle, 2016. Utilitarian preferences with multiple priors. Econometrica 84, 1181–1201.

Alon, Shiri, Schmeidler, David, 2014. Purely subjective maxmin expected utility. J. Econ. Theory 152, 382–412.

Anantanasuwong, Kanin, Kouwenberg, Roy, Mitchell, Olivia S., Peijnenburg, Kim, 2020. Ambiguity attitudes about investments: evidence from the field. Working paper.

Aryal, Gaurab, Stauber, Ronald, 2014. Trembles in extensive games with ambiguity averse players. Econ. Theory 57, 1–40.

Bade, Sophie, 2015. Randomization devices and the elicitation of ambiguity-averse preferences. J. Econ. Theory 159, 221–235.

Baillon, Aurélien, Bleichrodt, Han, 2015. Testing ambiguity models through the measurement of probabilities for gains and losses. Am. Econ. J. Microecon. 7, 77–100.

Baillon, Aurélien, Halevy, Yoram, Li, Chen, 2020. Experimental elicitation of ambiguity attitude using the random incentive system. Working paper.

Baillon, Aurélien, Huang, Zhenxing, Selim, Asli, Wakker, Peter P., 2018. Measuring ambiguity attitudes for all (natural) events. Econometrica 86, 1839–1858.

Binmore, Ken, Stewart, Lisa, Voorhoeve, Alex, 2012. How much ambiguity aversion? Finding indifferences between Ellsberg's risky and ambiguous bets. J. Risk Uncertain. 45, 215–238.

Brenner, Menachem, Izhakian, Yehuda, 2018. Asset pricing and ambiguity: empirical evidence. J. Financ. Econ. 130, 503–531.

Cappelli, Veronica Roberta, Cerreia-Vioglio, Simone, Maccheroni, Fabio, Marinacci, Massimo, Minardi, Stefania, 2020. Sources of uncertainty and subjective prices. J. Eur. Econ. Assoc., in press.

Cerreia-Vioglio, Simone, Dillenberger, David, Ortoleva, Pietro, Riella, Gil, 2019a. Deliberately stochastic. Am. Econ. Rev. 109, 2425–2445.

Cerreia-Vioglio, Simone, Maccheroni, Fabio, Marinacci, Massimo, 2019b. Ambiguity aversion and wealth effects. J. Econ. Theory, forthcoming.

Chateauneuf, Alain, Faro, José H., 2009. Ambiguity through confidence functions. J. Math. Econ. 45, 535–558.

Chateauneuf, Alain, Eichberger, Jürgen, Grant, Simon, 2007. Choice under uncertainty with the best and worst in mind: neo-additive capacities. J. Econ. Theory 137, 538–567.

Chew, Soo Hong, Li, King King, Chark, Robin, Zhong, Songfa, 2008. Source preference and ambiguity aversion: models and evidence from behavioral and neuroimaging experiments. In: Houser, Daniel, McCabe, Kevin (Eds.), Neuroeconomics (Advances in Health Economics and Health Services Research), vol. 20. JAI Press, Bingley, UK, pp. 179–201.

Dean, Mark, Ortoleva, Pietro, 2017. Allais, Ellsberg, and preferences for hedging. Theor. Econ. 12, 377–424.

Dimmock, Stephen G., Kouwenberg, Roy, Mitchell, Olivia S., Peijnenburg, Kim, 2015. Estimating ambiguity preferences and perceptions in multiple prior models: evidence from the field. J. Risk Uncertain. 51, 219–244.

Dimmock, Stephen G., Kouwenberg, Roy, Wakker, Peter P., 2016. Ambiguity attitudes in a large representative sample. Manag. Sci. 62, 1363–1380.

Dobbs, Ian M., 1991. A Bayesian approach to decision-making under ambiguity. Economica 58, 417–440.

Dow, James, Werlang, Sérgio R.C., 1992. Uncertainty aversion, risk aversion and the optimal choice of portfolio. Econometrica 60, 197–204.

Eichberger, Jürgen, Kelsey, David, 2011. Are the treasures of game theory ambiguous? Econ. Theory 48, 313–393.

Eichberger, Jürgen, Grant, Simon, Lefort, Jean-Philippe, 2012. Generalized neo-additive capacities and updating. Int. J. Econ. Theory 8, 237–257.

Einhorn, Hillel J., Hogarth, Robin M., 1985. Ambiguity and uncertainty in probabilistic inference. Psychol. Rev. 92, 433–461.

Eliaz, Kfir, Ortoleva, Pietro, 2016. Multidimensional ellsberg. Manag. Sci. 62, 2179–2197.

Ellsberg, Daniel, 1961. Risk, ambiguity and the Savage axioms. Q. J. Econ. 75, 643–669.

Epstein, Larry G., Schneider, Martin, 2010. Ambiguity and asset markets. Annu. Rev. Financ. Econ. 2, 315–346.

Fehr-Duda, Helga, Epper, Thomas, 2012. Probability and risk: foundations and economic implications of probability-dependent risk preferences. Annu. Rev. Econ. 4, 567–593.

Georgalos, Konstantinos, 2019. An experimental test of the predictive power of dynamic ambiguity models. J. Risk Uncertain. 59, 51–83.

Ghirardato, Paolo, Maccheroni, Fabio, Marinacci, Massimo, 2004. Differentiating ambiguity and ambiguity attitude. J. Econ. Theory 118, 133–173.

Ghirardato, Paolo, Marinacci, Massimo, 2001. Risk, ambiguity, and the separation of utility and beliefs. Math. Oper. Res. 26, 864–890.

Gilboa, Itzhak, Marinacci, Massimo, 2016. Ambiguity and the Bayesian paradigm. In: Arló-Costa, Horacio, Hendricks, Vincent F., van Benthem, Johan F.A.K. (Eds.), Readings in Formal Epistemology. Springer, Berlin, pp. 385–439.

Goeree, Jacob K., Holt, Charles A., 2001. Ten little treasures of game theory and ten intuitive contradictions. Am. Econ. Rev. 91, 1402–1422.

Grant, Simon, Polak, Ben, 2013. Mean-dispersion preferences and constant absolute uncertainty aversion. J. Econ. Theory 148, 1361–1398.

Gul, Faruk, Pesendorfer, Wolfgang, 2014. Expected uncertain utility theory. Econometrica 82, 1–39.

Gul, Faruk, Pesendorfer, Wolfgang, 2015. Hurwicz expected utility and subjective sources. J. Econ. Theory 159, 465–488.

Gul, Faruk, Pesendorfer, Wolfgang, 2020. Calibrated uncertainty. J. Econ. Theory 188, 105016.

Hansen, Lars P., Sargent, Thomas J., 2001. Robust control and model uncertainty. Am. Econ. Rev. Pap. Proc. 91, 60–66.

Harvey, Charles M., 1990. Structural prescriptive models of risk attitude. Manag. Sci. 36, 1479–1501.

Hill, R. Carter, Griffiths, William E., Lim, Guay C., 2008. Principles of Econometrics, 3rd edn. Wiley, New York.

Hey, John D., Lotito, Gianna, Maffioletti, Anna, 2010. The descriptive and predictive adequacy of theories of decision making under uncertainty/ambiguity. J. Risk Uncertain. 41, 81–111.

Hodges, Joseph L., Lehmann, Erich L., 1952. The use of previous experience in reaching statistical decisions. Ann. Math. Stat. 23, 396–407.

Hull, John C., 2017. Options, Futures, and Other Derivatives, 10th edn. Englewood Cliffs, Prentice-Hall, NJ.

Izhakian, Yehuda, 2017. Expected utility with uncertain probabilities theory. J. Math. Econ. 69, 91–103.

Izhakian, Yehuda, Brenner, Menachem, 2011. The uncertainty premium in an ambiguous economy. Q. J. Finance 1, 323–354.

Johnson, Cathleen, Baillon, Aurélien, Bleichrodt, Han, Li, Zhihua, van Dolder, Dennie, Wakker, Peter P., 2021. Prince: an improved method for measuring incentivized preferences. J. Risk Uncertain. 62, 1–28.

Kahn, Barbara E., Sarin, Rakesh K., 1988. Modeling ambiguity in decisions under uncertainty. J. Consum. Res. 15, 265–272.

Kahneman, Daniel, Tversky, Amos, 1979. Prospect theory: an analysis of decision under risk. Econometrica 47, 263–291.

Kilka, Michael, Weber, Martin, 2001. What determines the shape of the probability weighting function under uncertainty. Manag. Sci. 47, 1712–1726.

Klibanoff, Peter, Marinacci, Massimo, Mukerji, Sujoy, 2005. A smooth model of decision making under ambiguity. Econometrica 73, 1849–1892.

Kocher, Martin G., Lahno, Amrei Marie, Trautmann, Stefan T., 2018. Ambiguity aversion is not universal. Eur. Econ. Rev. 101, 268–283.

l'Haridon, Olivier, Vieider, Ferdinand, Aycinena, Diego, Bandur, Augustinus, Belianin, Alexis, Cingl, Lubomir, Kothiyal, Amit, Martinsson, Peter, 2018. Off the charts: massive unexplained heterogeneity in a global study of ambiguity attitudes. Rev. Econ. Stat. 100, 664–677.

Lehrer, Ehud, Teper, Roee, 2015. Subjective independence and concave expected utility. J. Econ. Theory 158, 33–53.

Maccheroni, Fabio, Marinacci, Massimo, Ruffino, Doriana, 2013. Alpha as ambiguity: robust mean-variance portfolio analysis. Econometrica 81, 1075–1113.

Maccheroni, Fabio, Marinacci, Massimo, Rustichini, Aldo, 2006. Ambiguity aversion, robustness, and the variational representation of preferences. Econometrica 74, 1447–1498.

Machina, Mark J., 2011. Event-separability in the ellsberg urn. Econ. Theory 48, 425–436.

Machina, Mark J., Schmeidler, David, 1992. A more robust definition of subjective probability. Econometrica 60, 745–780.

Mackenzie, Andrew, 2019. A foundation for probabilistic beliefs with or without atoms. Theor. Econ. 14, 709–778.

Neilson, William S., 2010. A simplified axiomatic approach to ambiguity aversion. J. Risk Uncertain. 41, 113–124.

Oechssler, Jörg, Rau, Hannes, Roomets, Alex, 2019. Hedging, ambiguity, and the reversal of order axiom. Games Econ. Behav. 117, 380–387.

Ortoleva, Pietro, 2012. Modeling the change of paradigm: non-Bayesian reactions to unexpected news. Am. Econ. Rev. 102, 2410–2436.

Pratt, John W., 1964. Risk aversion in the small and in the large. Econometrica 32, 122–136.

Savage, Leonard J., 1954. The Foundations of Statistics. Wiley, New York, 2nd edn. Dover Publications, New York, 1972.

Schmeidler, David, 1989. Subjective probability and expected utility without additivity. Econometrica 57, 571–587.

Siniscalchi, Marciano, 2009. Vector expected utility and attitudes toward variation. Econometrica 77, 801–855.

Skiadas, Costis, 2013. Scale-invariant uncertainty-averse preferences and source-dependent constant relative risk aversion. Theor. Econ. 8, 59–93.

Starmer, Chris, Sugden, Robert, 1991. Does the random-lottery incentive system elicit true preferences? An experimental investigation. Am. Econ. Rev. 81, 971–978.

Stock, James H., Watson, Mark W., 2015. Introduction to Econometrics. Pearson Education, Reading, Mass.

Trautmann, Stefan T., van de Kuilen, Gijs, 2015. Ambiguity attitudes. In: Keren, Gideon, Wu, George (Eds.), The Wiley Blackwell Handbook of Judgment and Decision Making (Ch. 3). Blackwell, Oxford, UK, pp. 89–116.

Tversky, Amos, Fox, Craig R., 1995. Weighing risk and uncertainty. Psychol. Rev. 102, 269–283.

Tversky, Amos, Wakker, Peter P., 1995. Risk attitudes and decision weights. Econometrica 63, 1255–1280.

Villegas, Cesáreo, 1964. On quantitative probability σ-algebras. Ann. Math. Stat. 35, 1787–1796.

Viscusi, W. Kip, Magat, Wesley A., 1992. Bayesian decisions with ambiguous belief aversion. J. Risk Uncertain. 5, 371–387.

Wakker, Peter P., 2010. Prospect Theory: For Risk and Ambiguity. Cambridge University Press, Cambridge, UK.

Walley, Peter, 1991. Statistical Reasoning with Imprecise Probabilities. Chapman and Hall, London.

Wu, George, Gonzalez, Richard, 1999. Nonlinear decision weights in choice under uncertainty. Manag. Sci. 45, 74–85.