Belief hedges: applying ambiguity measurements to all events and all ambiguity models[1]

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

a: Erasmus School of Economics, Erasmus University Rotterdam, Rotterdam, the Netherlands, Baillon@ese.eur.nl, Bleichrodt@ese.eur.nl, c.li@ese.eur.nl; Wakker@ese.eur.nl

* corresponding author

November, 2020

Abstract

We introduce belief hedges, which combine a properly chosen set of events whose uncertain subjective beliefs neutralize each other. They protect against unknown probabilities in ambiguity measurements, resulting in ambiguity indexes that improve virtually all preceding ones. Unlike predecessors that were only shown to be valid under one theory, our indexes are valid under all popular ambiguity theories. This frees practitioners from the choice overload given the many existing theories. Our indexes are directly observable for application-relevant events, with increased descriptive validity by not requiring expected utility for risk or two-stage optimization. Belief hedges make ambiguity theories widely applicable.

JEL-Classification: D81, C91

Keywords: Subjective beliefs; Ambiguity aversion; Ellsberg paradox; Sources of uncertainty

---

[1] A first draft of this paper was entitled: "Balanced Design: The Key to Measuring Ambiguity Attitudes when Beliefs Are Unknown."

1. Introduction

Hedging is a central concept in finance. Uncertain monetary outcomes ("gambles") are turned into certainties without using any further information about the relevant uncertainties. A hedge, instead, combines a properly chosen set of gambles so that their uncertainties neutralize each other, whatever they are. This paper introduces an analog for subjective beliefs, called belief hedges. A well-known problem in the measurement of ambiguity[2] attitudes is that there is uncertainty about the subjective beliefs of a decision maker, and these beliefs may confound the measurement. A belief hedge combines a properly chosen set of events so that their uncertain subjective beliefs neutralize each other, whatever they are. Thus, we can measure ambiguity attitudes without needing any further information about the relevant beliefs. We can now directly handle real-life uncertainties that are relevant for applications, with no more need to resort to artificial events such as Ellsberg urns or experimenter-specified probability intervals. This increases external validity, and the motivation of clients and subjects. Using our belief hedges, we will introduce general indexes of ambiguity aversion and insensitivity.

In an experiment on time pressure, Baillon et al. (2018) were the first to measure ambiguity attitudes without needing information about subjective beliefs. Their primary aim was to simplify Abdellaoui et al's (2011) source method. Their approach only works for three-fold partitions, and no (theoretical) justification was given. We show that their domain is a special case of a belief hedge. This provides a theoretical justification for their approach. We remove their restriction of three-fold uncertainty, providing the warranted flexibility for applications (Examples 4, 12, and 19). We identify the relevant general concepts (belief hedges and indexes) through necessary and sufficient axioms, and show that these concepts are justified by underlying econometric principles. Section 7 gives further details on our contribution to Baillon et al. (2018).

---

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

We show that our indexes do not only generalize those of Baillon et al., but most other indexes proposed in the literature as yet.[3] Thus, our indexes are not only valid under the source method, but also under many existing ambiguity theories. This is desirable given that there are over a dozen of ambiguity models.[4] Unlike their predecessors, our indexes do not need expected utility for risk, or two-stage stimuli and dynamic decision principles, making them descriptively valid and tractable. Our indexes operationalize existing ones in the following sense. Most existing indexes use theoretical constructs[5] such as nonadditive measures or sets of priors in their definition. This means that they are only indirectly observable. Our indexes can directly be elicited from preferences, with no need for data fitting or the additional assumptions it requires (choice of an error model and of a parametric specification). They thus show, in particular, how the indexes in the literature can be made directly observable.

A detailed outline is as follows. Unlike preceding papers, the first part of this paper does not commit to any ambiguity model when introducing our indexes. This absence underscores that our indexes do not involve parametric fittings, and can directly be revealed from preferences. In the absence of a decision model, the first part of our paper can only provide intuitive plausibility arguments. We use econometric concepts to do so. Following basic definitions (§2.1), §2.2 introduces belief hedges for ambiguity aversion, and a corresponding aversion index—an average ambiguity premium. This section, while elementary, already conveys the main novelty of belief hedges, explaining why they make artificial ambiguities (e.g. Ellsberg urns) redundant. Section 2.3 presents the same result for a second index of ambiguity, capturing insensitivity. This index is mathematically more complex but empirically useful. It reflects *changes* in aversion rather than aversion itself. It captures, for instance, by how much the decision maker underestimates the marginal benefits of prevention.

---

[3] Baillon et al. (2018) referred forward to the present paper for this result.

[4] References are in Online Appendix OA.1. Throughout this paper, OA.x refers to a literature survey in §x in Online Appendix OA.

[5] Theoretical constructs are not directly observable, but derive their empirical meaning indirectly in combination with other theoretical constructs (such as utility), and only within some assumed model (Cozic and Hill 2015).

Section 2.4 shows theoretically that our two indexes properly reflect aversion and insensitivity, and that the indexes concern distinct, orthogonal, components in the variance in our data. Section 3 gives tractable examples of belief hedges and a preference foundation of our indexes. It shows that belief hedges provide practitioners with the required flexibility and feasibility to tackle complex empirical problems.

The second part of this paper, starting in §4, does consider various ambiguity models. It first presents an extension of our indexes to many outcomes. In general, our indexes may be outcome dependent. They can serve as useful tools to examine this dependence in models where ambiguity attitudes indeed depend on outcomes, such as the smooth model. Section 5 considers the special case of outcome independence, which includes many existing ambiguity models such as biseparable utility, rank-dependent/Choquet expected utility, prospect theory, and multiple priors.

Section 6 shows that our indexes generalize and unify most indexes proposed before. For example, Dow and Werlang (1992) and Schmeidler (1989) proposed an index of ambiguity aversion using their nonadditive weighting function in Choquet expected utility, and assuming expected utility for risk. Under their assumptions, their weighting functions coincide with our matching probabilities and their index coincides with our aversion index. Our index remains valid, though, if expected utility for risk is violated (which is desirable for empirical purposes; Bruhin, Fehr-Duda, and Epper 2010), and also for ambiguity models other than Choquet expected utility. Our indexes also agree with the common qualitative orderings of ambiguity attitudes proposed in the literature, such as being more ambiguity averse. Because of the compatibility of our indexes with existing indexes and orderings, the arguments advanced in the literature for those indexes and orderings support our indexes. This gives a broad theoretical support to the intuitive claims made in the first part of this paper. It shows that our indexes capture the proper general components of ambiguity attitudes. The main mathematical surprise in this study was to find that most existing ambiguity indexes and orderings can be based on the same underlying econometric principles, explained in §3. This ensures conceptual soundness and good statistical performance.

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*, finite or infinite. Outcomes can be money amounts, health states, and so on. *Acts* map S to X and are finite-valued.[6] 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 formally define, or experimentally measure, the ambiguity indexes. A central question in our analysis will be which designs are suited for this purpose. In most of this paper (except §3) $\mathcal{H}$ is fixed, and then dependencies on it need not be expressed in 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, defined later, 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 analyzes

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

to what extent our indexes depend on the design (and, thus, on ν). Under empirically plausible assumptions the indexes are largely independent. Limitations will be discussed.

Throughout this paper, statistics refer to $\mathcal{H}$, with the following notation for functions $F, G: \mathcal{H} \to \Re$. $E(F) = \overline{F} = \frac{\sum_{E \in \mathcal{H}} F(E)}{|\mathcal{H}|}$ denotes average; $Var(F)$ denotes variance; $Cov(F, G)$ denotes covariance of F and G. Thus, all these statistics are to be taken over $\mathcal{H}$. 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.

2.2. Belief hedging for ambiguity aversion

Sections 2 and 3 introduce and analyze our concepts under minimal assumptions. Thus, these sections assume:

ASSUMPTION 1 [Two outcomes]. $X = \{\gamma, \theta\}$, with $\gamma \succ \theta$. □

Dimmock, Kouwenberg, and Wakker (2016 Theorem 3.1) showed that matching probabilities (defined next) 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. Our indexes will also be based on them. We thus assume that a *matching probability* $m(E)$ exists for every event E, defined by

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

Monotonicity implies that m is unique. Ambiguity reflects how $m(\cdot)$ deviates from a probability measure. For example, ambiguity aversion will imply $m(E) + m(E^c) < 1$, 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. □

In this and the next section we use a general framework where, besides Assumption 1, only Assumption 2 is made. Hence, the results here hold for virtually all ambiguity models considered today.

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

Ambiguity neutrality effectively means that subjective beliefs are treated the same way as objective beliefs. Observation 17 will show that the condition for two fixed outcomes as defined here implies general ambiguity neutrality under many ambiguity models.Ambiguity neutrality is violated in the following example. We chose it because it refers to an existing experiment. As a price to pay, it involves some game-theoretic details.

EXAMPLE 4 [Running Example]. The decision maker plays the following two-player minimum effort coordination game from Goeree and Holt (2001) in the version analyzed theoretically by Eichberger and Kelsey (2011). The players have to simultaneously and independently choose an effort level from $\{115, 125, 135, 145, 155, 165\}$, being act f or f_j for the decision maker and state s or s_j for her opponent, at marginal cost $c = 0.9$. They then receive the outcome $\min\{f, s\} - c \times e$, where e denotes own effort level. Optimal for the decision maker is to choose the same effort level as her opponent, but she does not know this level.

In our imaginary variation, we assume that we also observe the decision maker's matching probabilities through the following indifferences:[7]

$15_{s_1}0 \sim 15_{0.50}0$; that is, $m(s_1) = 0.50$;
$15_{s_2}0 \sim 15_{s_3}0 \sim 15_{s_4}0 \sim 15_{s_5}0 \sim 15_{0.20}0$; that is, $m(s_2) = \cdots = m(s_5) = 0.20$;
$15_{s_6}0 \sim 15_{0.30}0$; that is, $m(s_6) = 0.30$;
$0_{s_1}15 \sim 0_{55}15$; that is, $m({s_1}^c) = 0.45$;
$0_{s_2}15 \sim 0_{s_3}15 \sim 0_{s_4}15 \sim 0_{s_5}15 \sim 0_{0.30}15$; that is, $m({s_2}^c) = \cdots = m({s_5}^c) = 0.70$;

[7] Such side 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 side measurement (Chierchia, Nagel, and Coricelli 2018).

$0_{s_6}15 \sim 0_{0.35}15$; that is, $m({s_6}^c) = 0.65$.

The decision maker violates ambiguity neutrality as m is not a probability measure, violating additivity. For instance, $m(s_1) + m({s_1}^c) = 0.95 < 1$.

We will illustrate the techniques of our paper for this example in what follows, and summarize and discuss the results in Example 19. □

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*.[8] These differences reflect an ambiguity premium, i.e., willingness to pay—in probability (belief) units—to avoid ambiguity. The bigger the aversion, the bigger this premium (Viscusi and Magat 1992; Dimmock, Kouwenberg, and Wakker 2016). Ideally, with some observations $P(E) - m(E)$ available for a number of events E, we would like to define our aversion index as the average level of differences

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

The aforementioned references considered Ellsberg urns, where P could be derived from symmetry assumptions. For natural events the problem is that we do not know the a-neutral P. In Example 4, it need not agree with the actual choice percentages in Table 1, as those were unknown to the decision maker.

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

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

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

[8] They can be interpreted as the beliefs of the ambiguity neutral twin of the decision maker, i.e., the beliefs if the decision maker 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:

each state s appears in exactly half of the events in $\mathcal{H}$. (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.[9] 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

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

This way, by using l-hedging, we have captured Eq. 2 without needing to know P. The index reflects how much success probability one is willing to give up to avoid ambiguity. In the Anscombe-Aumann framework (expected utility) this reflects the proportion of success-utility one is willing to pay. For moderate stakes and approximately linear utility, b then is the proportion of the gain one is willing to pay to avoid ambiguity. In Example 4, we have the data for $\mathcal{H}$ containing all singletons and their complements, which satisfies l-hedging. We get $b = 0.08$, suggesting weak ambiguity aversion. Note that, because of l-hedging, this is the average ambiguity premium $\overline{P - m}$ for every probability measure P on S, and this is why we need not know P.

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

The analysis in this section was simple from an algebraic perspective. But because of it, there is no more need to rely on Ellsbergian informational symmetry of the events to measure ambiguity aversion, once we have ensured l-belief hedges through the measurement design $\mathcal{H}$, and we can directly use the application-relevant events. This will increase both validity and the motivation of subjects and clients.

2.3. Belief hedging for insensitivity

Theoretically and normatively motivated ambiguity models have focused on ambiguity aversion, a motivational component of ambiguity attitude.However, recent empirical studies have found richer phenomena (Anantanasuwong et al. 2020; l'Haridon et al. 2018; Kocher, Lahno, and Trautmann 2018). Whereas for likely events there indeed is strong ambiguity aversion, it gets weaker for events of moderate likelihood, and for low likelihood events it reverses. Then ambiguity aversion even predicts in the wrong direction (Trautmann and van de Kuilen 2015).[10]

The aforementioned likelihood dependence shows a general drift towards fifty-fifty, with insufficient discriminatory power and insufficient responsiveness toward belief changes in the middle region. It is a similar kind of insensitivity as exhibited by inverse-S probability weighting for risk, where weights in the middle are also moved toward fifty-fifty (Fehr-Duda and Epper 2012), but now it concerns ambiguity attitudes. We incorporate this effect as a second component of ambiguity attitude, and interpret it to be cognitive, an interpretation supported by Anantanasuwong et al.'s (2020) large-scale empirical study of a representative sample of financial investors. This component reflects lack of understanding of ambiguity, which comes prior to any aversion or seeking. The decision maker takes ambiguous events (too much) as one blur. We use the term a(mbiguity-generated) insensitivity to refer to the insensitivity generated by ambiguity.

This section explains our measurement of the second component, again independently from beliefs, which is similar to the preceding section but mathematically more involved. Ideally, we would like to use the most common measure of responsiveness of m with respect to P, being the sensitivity

[10] 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.

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

This index has been widely used as the slope in regressions, and as β in the CAPM model in finance. 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 priors models, our insensitivity index coincides with the size of the set of priors (§6.3). In general, the larger the set of priors (perception of ambiguity), the more events are treated alike, as one blur, 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 the same way, with all $\beta_E\alpha$ indifferent and with, indeed, maximal insensitivity. Section 6 provides similar results for other popular ambiguity theories, where insensitivity is often interpreted as perception of ambiguity. Our insensitivity index shows a way to directly measure this based on revealed preferences.

To measure our insensitivity index, the problem we face is, again, that the a-neutral P is unknown. Our solution is to ensure that we can replace P by the event size ν (as if P were uniform over atoms), irrespective of what P is (as long as it is plausible). The intuition is to ensure that the event size properly reflects the average probability P (over events of the same size) in the sense that they perfectly co-vary with each other. The following condition is necessary and sufficient for this purpose.

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

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

This condition requires that the total size of events containing each fixed state s (i.e., containing each atom E_i, through any $s \in E_i$), is a constant. One can verify that 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 here, leaving the derivation to Appendix A—all proofs in this paper are given in the appendix.

Intuitively, v-hedging ensures that each state s and, hence, each atom E_i (through any $s \in E_i$) appears equally often in big events and, accordingly (by Eq. 4), in small events. Therefore, shifting belief between states/events in S, does not lead to more total belief for big events (or, correspondingly, to less total belief for small events) in $\mathcal{H}$. In Example 4, one can verify that the total belief for all small events of size $1/6$ is always 1 and the total belief for all big events of size $5/6$ is always 5, no matter what P is. Our conditions ensure that the total belief is the same for each belief, including ν. Hence, the extent to which the sum of m over- or underweights big or small events in $\mathcal{H}$ cannot be due to beliefs, and it must reflect attitude. That beliefs do not matter and may as well be assumed to be ν justifies the following approximation of sensitivity of m with respect to P by sensitivity of m with respect to ν. Appendix A shows that Eq. 8 provides a good first-order approximation under common econometric assumptions. Proposition 21 there will derive exact equality in some cases, implying good fit in all empirically plausible cases.

$$\frac{Cov(m,P)}{Var(P)} \approx \frac{Cov(m,\nu)}{Var(\nu)}. \tag{8}$$

For Eq. 8, we need one more assumption, to avoid degeneracy:

ASSUMPTION 8 [nondegeneracy]. $\mathcal{H}$ does not contain $\emptyset$ or S. All atoms E_j are nonnull. ν is not constant on $\mathcal{H}$. □

Regarding the first part of the assumption, insensitivity concerns intermediate events away from the extremes and, hence, we exclude the extreme events.[11] This entails no loss of information because the m values of $\emptyset$ and S are 0 and 1, respectively, by monotonicity. As for the second part of the Assumption, null events do not affect preference and, hence, can be made to disappear 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. For the final part, because we derive insensitivity from variations in event size, we need event size to be nonconstant—also after excluding $\emptyset$ and S. This implies $n \geq 3$. The condition ensures that $Var(\nu)$ is positive,

[11] 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 not using events in the measurement design with a-neutral probabilities below 0.05 or above 0.95.

so that the ratios are well-defined. $\mathcal{H}$ is a *belief hedge*, or *hedge* for short, if both l-hedging and v-hedging hold.

DEFINITION 9. If Assumption 8 and belief hedging (Eqs. 4 and 7) hold, then the *index of a(mbiguity-generated) insensitivity* is

$$a = 1 - \frac{Cov(m,v)}{Var(v)} . \tag{9}$$

Recall that Cov and Var concern variation within $\mathcal{H}$ and can be calculated exactly from the matching probability data collected for all events in $\mathcal{H}$. Eq. 10 below gives a simple special case of Eq. 9 that can readily be calculated using paper and pencil. Whereas the aversion parameter captures how much probability is lost due to ambiguity, the insensitivity index captures the part of *changes* in probability lost due to ambiguity (away from the extreme likelihoods). It thus captures the degree of underreaction to new information, and is relevant, for instance, in evaluations of precautionary measures. An index $\text{a} = 0.43$ (as in Example 4) means that the decision maker underestimates the marginal benefits of precautionary measures by a factor of almost 2.

Throughout the rest of the paper we assume that Assumptions 2 and 8 hold, explicitly in theorems and implicitly elsewhere. We, finally, return to our aversion index. The beginning of this section indicated that the prevailing empirical finding for unlikely events is ambiguity seeking. Hence, had we mainly used unlikely events in $\mathcal{H}$, e.g. the singletons in Example 4, then the average $\overline{P - m}$ would have been small or even negative, underestimating ambiguity aversion. Using mainly likely events in $\mathcal{H}$ would overestimate ambiguity aversion. L-hedging has avoided such biases by taking average event-size 1/2.

2.4. Theoretical justifications of belief hedges

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

THEOREM 10. Under Assumptions 1, 2, and 8, $b = a = 0$ for all ambiguity neutral decision makers if and only if the measurement design $\mathcal{H}$ is a belief hedge. Then the supremum value of both b and a is 1.[12] □

Psychologically, we interpret aversion and insensitivity as two distinct components of ambiguity attitudes. This interpretation is supported by an orthogonality—in the usual Euclidean sense as used for instance in statistical analyses of variance (Appendix B)—of the indexes. Although conceptual orthogonality need not preclude any empirical relation, Anantanasuwong et al. (2020) found orthogonality empirically in this case.

THEOREM 11. Under Assumptions 1, 2, and 8, the indexes a and b capture orthogonal components of the variance of the data. □

3. Which design to use and a preference foundation

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 $\overline{m_c} = \frac{m(E_1 \cup E_2)+m(E_1 \cup E_3)+(E_2 \cup E_3)}{3}$. They used the following definitions which, by substitution (Online Appendix OB) are identical to ours:

$$b = 1 - \overline{m_c} - \overline{m_s} \text{ and } a = 3\left(\frac{1}{3} - (\overline{m_c} - \overline{m_s})\right). \tag{10}$$

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 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 that event with E_i replaced by E_j is also

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

contained in $\mathcal{H}$. This is satisfied if $\mathcal{H}$ is the *full design*, i.e., contains all unions of E_js except S and $\emptyset$, as with $\mathcal{H}\{E_1, E_2, E_3\}$. It is also satisfied if $\mathcal{H}$ is the *basic design*, i.e., contains all one-atom events and their complements. Further, disjoint unions of belief hedges are always belief hedges again.

The examples show that there is much flexibility in belief hedges. The smallest one possible results from a three-fold partition of S and then a full design, being the design of Baillon et al. (2018). Simplicity goes at the cost of reliability though, and the richer $\mathcal{H}$ is, the more reliable and valid the estimated indexes will be. As default we recommend a basic design with a partition of S that specifies all relevant uncertainties[13], such as the six possible effort levels of the opponent in Example 4. This design involves all relevant atoms, considers likely and unlikely events (where ambiguity is strongest), and grows linearly with the number of atoms so that it is tractable. A big pro of the basic design, as well as richer designs, is that we get enough equalities to also estimate subjective beliefs (Example 19). In many situations, especially if analyzing real-world data sets, one may not have much control over the data received, and then the flexibility of general belief hedges is useful.

In the basic design of Example 4 one could, at will, add $\{\{s_1, s_2, s_3\}, \{s_4, s_5, s_6\}\}$ there, or any other pair of disjoint three-state events. Such additions are desirable if some such events are of special relevance, or are expected to show deviating behavior (see Example 12 below). One can further select events for being easy to relate to for the subjects and for avoiding biases. For the sake of brevity, this theoretical paper leaves experimental implementations, and applications to real-world data sets, to future studies.

In general, different designs need not give the same indexes, as the following example shows.

EXAMPLE 12. Consider an Ellsberg urn with 90 balls numbered 1-90, the first 30 red, the last 60 black or yellow in 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

[13] Preferably, this does not involve very many 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.

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 odd/even and color is needed to obtain the average aversion/insensitivity over all events. □

Ambiguity is too rich a domain to expect that one ambiguity attitude for a decision maker can cover all events. There can be many kinds of (source) preferences and (lacks of) understanding of uncertainty beyond risk, where emotions and confusions play a role beyond the degree to which probabilities are known or unknown (Tversky and Fox 1995). Ambiguity attitudes depend on sources of uncertainty similarly as utility functions depend on commodities (Cappelli et al. 2018). Our indexes can serve as tools to examine such dependencies and emotions, and this concerns a large and important topic for future research. We next investigate when different designs do give the same indexes.

DEFINITION 13. The indexes *perfectly fit* 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

$$P(E) = 0 \;\Rightarrow\; m(E) = 0;$$
$$0 < P(E) < 1 \;\Rightarrow\; m(E) = \tau + \sigma P(E);$$
$$P(E) = 1 \;\Rightarrow\; m(E) = 1. \tag{11}$$

We call W *neo-additive* if the three implications in Eq. 11 hold with W instead of m, where furthermore $\sigma > 0$ and all $P(E_i) > 0$ (to satisfy monotonicity[14]). Under Assumptions 1, 2, and 8, and neo-additivity of m (Eq. 11), substitution (Online Appendix OB) gives:

[14] This also avoids some open mathematical problems in Chateauneuf, Eichberger, and Grant (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.

$$b = 1 - 2\tau - \sigma \quad \text{and} \quad a = 1 - \sigma. \tag{12}$$

As is common in axiomatizations, we assume complete information about preferences. That is, we consider all measurement designs and use m for all events.[15] And, as common in axiomatizations, we assume a continuum domain (Dietrich 2018 pp. 18-19). We do so through the following conditions of Villegas (1964). We say that m is *fine*[16] if for each nonnull event A there exists an event $B \subset A$ such that $m(A) > m(B) > 0$. For any P, P is *fine* if the same holds for P instead of m. *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 14[17]. 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"[18]) 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, Grant, and Lefort 2012 p. 238 penultimate paragraph). Theorem 14 provides another one. In particular, it shows a new way to test the neo-additive model.

The rest of this section discusses (limitations of) applicability of our indexes. Theorem 14 is empirically reassuring because the neo-additive model performs well

---

[15] So far we, in fact, only used m on one fixed measurement design.

[16] 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).

[17] 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.

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

empirically which, together with its tractability, has made it popular. In particular, it captures the empirically prevailing four-fold pattern of ambiguity aversion (Trautmann and van de Kuilen 2015). If the neo-additive fit is not perfect, then our indexes can serve as pragmatic estimates, similarly as linear regressions and CRRA indexes are often used pragmatically. We next discuss similar limitations of our indexes.

Our pair of indexes surely cannot fit all data well if our source of uncertainty involves different uncertainty mechanisms, as in Example 12, where the source in this sense 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. We still recommend our pair of indexes there as the best ("average") summary using only one pair of numbers. Our indexes may not work well if very unlikely events are incorporated into the measurement design. Such events are known to involve many irregularities (Kahneman and Tversky 1979). We recommend avoiding them in applications, e.g., by imposing boundary conditions (Tversky and Wakker 1995; see Wakker's 2010 insensitivity region). Details on underlying econometric assumptions are discussed at the end of Appendix A. If practitioners reckon with the limitations just discussed, then our indexes can serve well to capture ambiguity attitudes.

Theorem 14 serves our indexes similarly as the classical result in expected utility for risk where the CRRA index perfectly captures the same (relative) risk aversion irrespective of the stimuli used if and only if utility is from the CRRA family (see, e.g., Theorem 3 in Harvey 1990). The CRRA index is tractable and performs well empirically, even though it shares similar limitations as our indexes. Empirically, the fit is usually not perfect as relative risk aversion is mostly not constant. In such situations, however, no CRRA index can fit all the data, but may still be the best ("average") summary. Also, it does not work well if extreme outcomes are incorporated into the measurement. Nevertheless, if practitioners reckon with these limitations, the CRRA index can serve well to capture risk attitudes.

Whereas, for instance, expected utility with CRRA utility is a one-parameter model, ambiguity models involve many additional parameters (including utility functions, nonexpected utility parameters of risk attitudes such as probability weighting, and subjective beliefs) besides the parameters of interest in this paper, capturing ambiguity. Whereas prior studies needed to jointly estimate all model parameters in order to obtain ambiguity measurement, we only need few indifferences. We do so without imposing any restriction on the additional parameters. Hence, we greatly simplify the empirical measurement by making many high-dimensional unknowns (utility, probability weighting, beliefs) drop from the equations. This is why a few indifferences suffice to give insights into a high-dimensional parametric model, without requiring complex data fittings.

4. Extension to many outcomes and outcome-dependent ambiguity models

The following sections drop Assumption 1 (two outcomes) and consider general outcome sets X. The results obtained before can trivially be extended as follows, without commitment to any ambiguity model.

OBSERVATION 15. All results of §2 and 3, including Theorems 10, 11, 14, and Eq. 10 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. □

In several models, ambiguity attitudes depend on the outcomes considered (Chew et al. 2008; see OA.2, i.e., Online Appendix OA.2). Then the indexes in Observation 15 will depend on the outcomes γ, θ chosen, and can be used to investigate this dependence. For example, constant ambiguity aversion w.r.t. absolute utility increments (Grant and Polak 2013), or w.r.t. proportional utility increments (Chateauneuf and Faro 2009), or these conditions w.r.t. wealth increments (Cerreia-Vioglio, Maccheroni, and Marinacci 2019), are inherited by matching probabilities and our indexes. These conditions can, therefore, be tested using our indexes.

The most popular outcome-dependent ambiguity model is the smooth model (Klibanoff, Marinacci, and Mukerji 2005). We obtain as aversion index (Online Appendix OC):

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

Here $A = -\frac{\varphi''}{\varphi'}$ is the Arrow-Pratt index of the function φ that captures ambiguity in the smooth model by transforming risky utility (Klibanoff, Marinacci, and Mukerji

2005 p. 1865), 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. 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. Eq. 13 makes the average ambiguity aversion of the smooth model, involving the not directly observable p, A, and σ^2, directly observable because b is.

We obtain as insensitivity index (Online Appendix OC):

$$a = \frac{1}{2}\frac{Cov(\sigma^2 A(p), v)}{Var(v)} + o(\sigma^2) . \qquad (14)$$

It captures how the aversion premium ($\sigma^2 A(p)$ as in Eq. 13) increases with event size v, which indeed reflects sensitivity. The ambiguity attitude analyzed here depends on the outcome interval $[\theta, \gamma]$ considered, as is typical of the smooth model (e.g., Klibanoff, Marinacci, and Mukerji 2005 Proposition 4).

We obtain outcome independence for the special case of $\varphi(x) = -e^{-\rho x}$ (Klibanoff, Marinacci, and Mukerji 2005 Proposition 2). Then:

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

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

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

5. Extension to outcome-independent ambiguity models

This section continues to consider general outcome sets X, dropping Assumption 1. Many models assume that ambiguity attitudes are outcome independent.[19] Then so

---

[19] They include biseparable utility (Ghirardato and Marinacci 2001) and, thus, Choquet expected utility or rank-dependent utility, prospect theory for gains, maxmin EU, and the α-maxmin model. Further

will our indexes be. Those models are all special cases of the following one. For simplicity, we assume the existence of a worst outcome.[20] *Uniseparable utility* holds if there exists a worst outcome θ ($\forall \gamma \in X: \gamma \succcurlyeq \theta;\ \exists \gamma \succ \theta$) such that

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

represents preferences for prospects with at most one outcome γ other than θ. For monetary gain outcomes, typically $\theta = 0$; in the health domain, often θ = death. Under prospect theory, θ is the reference outcome. U is the 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, $w: [0,1] \rightarrow [0,1]$ is a *(probability) weighting function*, with $w(0) = 0$, $w(1) = 1$, and w 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 16. $m(E) = p$ if $\gamma_E \theta \sim \gamma_p \theta$ for some $\gamma \succ \theta$. □

By Eq. 17, Definition 16 is equivalent to $\gamma_A \theta \sim \gamma_p \theta$ for all $\gamma \succ \theta$ (and it is equivalent to $W(A) = w(p)$). It, thus, extends our preceding definition (Eq. 1) to more than two outcomes.

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.

OBSERVATION 17. All results of §2 and §3 remain valid, including Theorems 10, 11, 14, and Eq. 10 if we replace Assumption 1 by uniseparable utility and use Definition

---

included are 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.

[20] We focus on gains so that sign-dependence, as in prospect theory for ambiguity, plays no role.

16 instead of Eq. 1. Ambiguity neutrality (Definition 3) then implies $W(.) = w\big(P(.)\big)$ for a subjective probability measure P (= m). □

The observation shows how our indexes and results can be applied to, basically, all event-driven ambiguity models. Observation 17 also shows that ambiguity neutrality in Definition 3, restricted to two fixed outcomes, agrees with common definitions. Ambiguity neutrality comprises both probabilistic sophistication (Machina and Schmeidler 1992) and indifference between subjective and objective probabilities (Dean and Ortoleva 2017 Footnote 31).

6. Generalizing and unifying existing ambiguity indexes and orderings

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

6.1. Qualitative ambiguity orderings

All papers that we are aware of (OA.5) define ambiguity neutrality as (a special case of) global probabilistic sophistication, sometimes as expected utility. Then m is an additive probability (Definition 3 and Observation 17) and both our indexes are 0 (Theorem 10), compatible with the existing definitions. The sign of b then properly reflects ambiguity aversion/seeking.

In virtually all papers in the literature, $\succcurlyeq^1$ is defined to be 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). The most general version is in Cerreia-Vioglio et al. (2011); further, see OA.6. This implies 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. In multiple priors models, set-inclusions of sets of priors have been considered (Ghirardato, Maccheroni, and Marinacci 2004 Proposition 6) that, for tractable subcases of multiple priors models, agree with our insensitivity index (Eq. 20 below). Tversky and Wakker (1995) considered comparative subadditivity for general weighting functions W. If applied to matching probabilities, they correspond with the indexes of Baillon and Bleichrodt (2015) (discussed in §6.2 below) and, therefore, this comparative subadditivity is compatible with our index a. Similarly, Tversky and Wakker's (1995) source preference conditions are compatible with b.

Some papers defined ambiguity indexes, and orderings, using premiums in monetary units rather than in our probability units (OA.7). 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.

6.2. Biseparable utility (including Choquet expected utility)

Many theories are special cases of uniseparable utility, including biseparable utility and Choquet expected utility, using 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 explicitly proposed by Dow and Werlang (1992). Commonly, expected utility is then assumed for risk, so that $m = W$, and we get:

OBSERVATION 18. 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[21] implies $b > 0$, ambiguity neutrality implies $b = 0$, and ambiguity seeking implies $b < 0$. □

Eq. 18 has been commonly used in theoretical studies (Klibanoff, Marinacci, and Mukerji 2005 Definition 7). MacCrimmon and Larsson (1979 p. 381-384) provided an

[21] 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.

early test. Because Eq. 18 uses the theoretical construct of W, originally it could not be readily implemented empirically. Our aversion index shows how to make it observable. Authors using Eq. 18 commonly assumed expected utility for risk, which our indexes do not need, increasing their descriptive validity.

Baillon and Bleichrodt (2015) considered a domain $\mathcal{H}(E_1, E_2, E_3)$ as in our Eq. 10, and measured five event-specific indexes. These indexes did not provide controls for beliefs. Our indexes show how their indexes can be aggregated to provide that control, capturing both aversion and insensitivity.[22] Our indexes are also compatible with those of Chateauneuf, Eichberger, and Grant (2007).[23] Under Choquet expected utility with expected utility for risk, our Theorem 14 provides an alternative axiomatization of Chateauneuf, Eichberger, and Grant's (2007) neo-additive model. Their model is in the intersection of Choquet expected utility and multiple priors models, to which we turn next.

6.3. Multiple priors

We consider some popular special cases of multiple prior models. Here, C denotes 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, Maccheroni, and Marinacci 2004), preferences maximize, for $\gamma \succcurlyeq \beta$:

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

with $W(E) = \alpha P_*(E) + (1 - \alpha)P^*(E)$ $(0 \le \alpha \le 1)$. Expected utility is assumed for risk. Maxmin expected utility is the special case of $\alpha = 1$Alon and Schmeidler 2014. We get, assuming complementation-closedness (v-hedging is not needed here):

---

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

[23] 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. 12 gives our indexes. Chateauneuf, Eichberger, and Grant (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.

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

Here $\left(\overline{P^* - P_*}\right)$ 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). Further, $2\alpha - 1$ (or, equivalently, α itself) is commonly taken as an index of ambiguity aversion. It is 0 under ambiguity neutrality. We discuss its relation with b after Eq. 20.

We next consider a tractable subclass of α-maxmin, the *ε-α-maxmin model* because here an index of insensitivity/perception has been proposed in the literature. Now $C = \{(1-\varepsilon)Q + \varepsilon T\}$, 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, Eichberger, and Grant 2007). It is a subclass of the ε-contamination model (Ellsberg 1961 pp. 663-669) that has been used in many fields (OA.3). Here ε $(= \overline{P^* - P_*})$, capturing the size of the set of priors, has been proposed as an index of ambiguity perception (Chateauneuf, Eichberger, and Grant 2007 p. 543; OA.4). Dimmock et al. (2015) showed:

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

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 ambiguity perception and aversion per unit of perception, capturing aversion in an absolute sense. 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.[24]

The following special case of α-maxmin was considered by Hey, Lotito, and Maffioletti (2010). 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 we have, with similar interpretations as before[25]:

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

---

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

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

Dimmock et al. (2015) showed how to make ambiguity aversion and the perceived level of ambiguity directly observable, without the need to measure utility U or the set of priors C, for Ellsberg urn events. Our contribution here is to abandon the restriction to Ellsberg urns (and expected utility for risk).

This subsection has shown how the currently popular indexes of ambiguity in multiple prior theories can be measured directly for application-relevant events. This can now be done with no more need to measure U or the set of priors C, or, as was common in applications hitherto, just ad hoc assuming a set C given exogenously.

6.4. The source method

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

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

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. 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, Kouwenberg, and Wakker (2016), abbreviated AD here, used the best neo-additive approximation of a function (w_S and $m(E)$, respectively) on the open interval $(0,1)$ by minimizing quadratic distance as in regular regressions. They then derived their indexes from this. We here do so for the function $m(E)$ (Eq. 11), where $\sigma \geq 0$ and τ are chosen to minimize distance.

AD needed a-neutral probabilities $P(E)$ specified beforehand, based on classical Ellsberg symmetry assumptions ($P = \nu$; Dimmock, Kouwenberg, and Wakker 2016) or on separate measurements (Abdellaoui et al. 2011). With τ and σ the best-fitting neo-additive parameters, AD defined (as in Eq. 12)

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

It is well-known from linear regression theory that then $a' = 1 - \frac{Cov(m,P)}{Var(P)}$, and that our index b always agrees with b'. By Eq. 8 and the results in Appendix A, our index a agrees well with a'.

Our contribution to the source method is that knowledge of P is no longer needed. Thus, the restriction to Ellsberg urns of Dimmock, Kouwenberg, and Wakker (2016) is no longer needed, and neither is the separate measurement of P (and U) in Abdellaoui et al. (2011).

7. Discussion

Whereas our indexes can be used beyond Ellsberg urns, it remains interesting to apply them to the widely studied Ellsberg urns. Example 12 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. Cappelli et al. (2018) give some theoretical suggestions. We next use our running example to illustrate some pros of our approach.

EXAMPLE 19 [Example 4 Continued]. Eichberger and Kelsey (2011), EK henceforth, also considered another marginal cost besides $c = 0.9$, being $c = 0.1$. The changes in choice percentages found (shifting towards high effort) were intuitive, but hard to explain by classical game theory. EK showed that ambiguity theories can give plausible explanations. For empirically plausible ambiguity attitudes they referred to another paper (Kilka and Weber 2001) that considered different uncertainties (and subjects) and inferred subjective beliefs from introspective judgments. Hence, it was not revealed-preference based. Their ranges of plausible ambiguity attitudes, reformulated here in terms of our indexes: $-0.15 \leq b \leq 0.12$ and $0.41 \leq a \leq 0.61$ (EK p. 319). These include the values ($b = 0.08$, $a = 0.43$) that we found in our imaginary example. Using belief hedges, we can measure ambiguity attitudes with the following pros: (1) they are of the players themselves; (2) they refer directly to the uncertainty relevant here (the effort level of the other player); (3) the subjective beliefs 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 pro 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. The contribution of this paper concerns the reversed direction: given the results of Baillon et al. (our Eq. 10), develop the general indexes and the concept of belief hedges. Finding Eq. 9 as the proper general concept of insensitivity, was the most challenging step in the project of these two papers. Validity of the general indexes was subsequently confirmed by theoretical justifications: preference axiomatizations and the common generalization of virtually all existing indexes, Baillon et al.'s included.[26] 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, 12, and 19). In many situations, especially when analyzing real-world data sets, one may not have much control over the data received, and then the flexibility of general belief hedges is desirable. Other practical pros over the special case of Baillon et al. (2018) were discussed at the beginning of §3. By the generalizations provided by this paper, ambiguity theories become widely applicable.

8. Conclusion

For a long time, ambiguity measurements were confined to artificial events such as secretized urns, because it was unknown how to control for unknown beliefs. We have introduced belief hedges for measuring ambiguity attitudes when subjective beliefs are unknown. Belief hedges extend the hedging concept from finance, where it provides protection against unknown outcomes, to ambiguity where it provides protection against unknown beliefs and, thus, the required controls. Through axiomatizations we identify belief hedges as necessary and sufficient for measuring ambiguity attitudes when beliefs are unknown. Thus, ambiguity attitudes can be directly measured for application-relevant events, and resorts to secretized urns are no longer needed.

Using belief hedges and some econometric concepts, we introduce two new indexes of ambiguity. They bring many improvements over existing indexes, and this

---

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

is the second contribution of this paper. Our indexes are general enough to provide the desired flexibility to practitioners (Examples 4, 12, and 19). They generalize indexes provided by Baillon et al. (2018) but, further, most other indexes proposed in the literature so far. They thus unify existing indexes, including ambiguity orderings. We show that those can all be based on some basic econometric principles, and our indexes are valid under virtually all existing ambiguity theories. Unlike their predecessors, they do not require expected utility for risk or multi-stage stimuli, which is desirable for empirical purposes. Further, they can accommodate ambiguity seeking for unlikely events which is, again, empirically desirable. And, unlike their predecessors in the literature, our indexes use no theoretical constructs. Hence, they can be directly revealed from preferences and in this sense operationalize the preceding indexes. In particular, our indexes need no measurements and data fittings of risk attitudes (utility/probability weighting) or a-neutral probabilities.

Appendix A. Goodness of fit of Eq. 8

Throughout this paper, for $F{:}\,\mathcal{H} \to \Re$, we write[27]: $\overline{F} = \frac{\sum_{E\in\mathcal{H}} F(E)}{|\mathcal{H}|}$; $Var(F) = \frac{\sum_{E\in\mathcal{H}}\left(F(E)-\overline{F}\right)^2}{|\mathcal{H}|}$; $Cov(F,G) = \frac{\sum_{E\in\mathcal{H}}\left(F(E)-\overline{F}\right)\times\left(G(E)-\overline{G}\right)}{|\mathcal{H}|}$.

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

LEMMA 20. Assume Assumption 8 and l-hedging. Equivalent are:

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

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:

$$\frac{Cov(1_s,\nu)}{Var(\nu)} = \frac{Cov\left(1_{E_i},\nu\right)}{Var(\nu)} = \frac{Cov(P,\nu)}{Var(\nu)} = \frac{Cov(\nu,\nu)}{Var(\nu)} = 1. \quad (25)$$

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

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.[28] The third equality follows because ν is a special case of a probability measure, and the last equality is by definition. □

We next turn to extraneous randomness in the dependency of m on P and ν. The above equality $\frac{Cov(P,\nu)}{Var(\nu)} = 1$ means that, on average, a change of one unit of ν

---

[27] 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.

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

generates one unit change of P. Hence, by Stock and Watson (2015 §12.1 and Eq. 12.7), Eq. 8 provides the best first-order approximation under common econometric assumptions together with the following critical assumption: m depends on ν 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 21. Under Assumptions 2 and 8 and belief hedging, Eq. 8 holds with exact equality if any of the following three conditions holds:

(i) $P = \nu$;
(ii) m is neo-additive;
(iii) P best fits m.[29]

PROOF. (i) is trivial, and (ii) follows from Eq. 12 (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 $\sum_{E \in \mathcal{H}} \big(m(E) - \tau - \sigma P(E)\big) = 0$. Thus, using Eq. 3,

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

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:

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

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

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

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

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

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. 8. □

Eq. 8 gives a good approximation if any of the three cases in Proposition 21 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. The data analysis in Online Appendix OD indeed found a good empirical fit. The average absolute discrepancy in Eq. 8 was 0.006. In 95% of the cases, the discrepancy was less than 0.01. The remaining 5% all concerned subjects who violated monotonicity (then our theoretical analysis makes no claims), with maximal discrepancy 0.27 for a highly erratic subject. We conclude that Eqs. 8 and 9 work well for all practical purposes.

Appendix B. Proofs except of Theorem 14

PROOF OF THEOREM 10. 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, $\frac{Cov(m,\nu)}{Var(\nu)} = 1$ 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 $\frac{Cov(1_s,\nu)}{Var(\nu)} = 1$ for all s which, by Eq. 24, implies v-hedging.

We, finally, turn to the supremum values of the indexes. b tends to its suppremum 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. □

PROOF OF THEOREM 11. 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* $\left(v(E) - \frac{1}{2}\right)_{E \in |\mathcal{H}|}$.[30] The aversion and insensitivity vectors are orthogonal because their inner product is $\sum \left(v(E) - \frac{1}{2}\right) = 0$. □

PROOF OF EQ. 19. Here we also assume complementation-closedness. (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, $m(E) + m(E^c) = \alpha P_*(E) + (1-\alpha)P^*(E) + \alpha(1 - P^*(E)) + (1-\alpha)(1 - P_*(E)) = 1 - (2\alpha - 1)\left(P^*(E) - P_*(E)\right)$. Finally, $= 1 - \overline{2m(E)} = 1 - \overline{m(E) - m(E^c)} = (2\alpha - 1)\left(\overline{P^* - P_*}\right)$. □

Appendix C. Proof of Theorem 14

That (i) implies (ii) in Theorem 14 follows because Eq. 12 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 12.

OBSERVATION 22. 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.

---

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

PROOF. The proof uses Lemmas 23-27.

LEMMA 23. 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 (Eq. 10) 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 (≈ additivity)—this is also the reason that we need two event continuity conditions, whereas for Villegas one is equivalent to the other.

LEMMA 24. 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 $\cup_{i=j}^{\infty} D_i$ 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 $S_\infty := D - \cup_{i=1}^{\infty} D_i$ 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, $\cup_{i=1}^{\infty} D_i = D$. By event continuity, $C := \cup_{i=1}^{J} D_j \succ B$ for J large enough. □

LEMMA 25. 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 24, 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 24, 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 23. QED

LEMMA 26. 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 23, $S - A - C \preccurlyeq S - B - C$. By Lemma 25, $A \cup C \succcurlyeq B \cup C$. The same reasoning holds with strict preferences. QED

Villegas used the following implication.

LEMMA 27. 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 26, and Villegas (1964, p. 1789 4th para). QED

Observation 22 now follows from Villegas (1964, Theorem 4.3). □

OBSERVATION 28. 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. 12). 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 $\frac{j}{3\times 2^{k+1}}$ (< ½) 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, Eq. 10 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 29. Perfect fit in Statement (ii) in Theorem 14 can be restricted to designs $\mathcal{H}\{E_1, E_2, E_3\}$. □

D.5. Further remarks (voor mezelf: over andere modellen)

ERASMUS SCHOOL OF ECONOMICS, ERASMUS UNIVERSITY ROTTERDAM

References

Abdellaoui, Mohammed, Aurélien Baillon, Laetitia Placido, & Peter P. Wakker (2011) "The Rich Domain of Uncertainty: Source Functions and Their Experimental Implementation," *American Economic Review* 101, 695–723.

Alon, Shiri & David Schmeidler (2014) "Purely Subjective Maxmin Expected Utility," *Journal of Economic Theory* 152, 382–412.

Anantanasuwong, Kanin, Roy Kouwenberg, Olivia S. Mitchell, & Kim Peijnenburg (2020) "Ambiguity Attitudes about Investments: Evidence from the Field," working paper.

Baillon, Aurélien & Han Bleichrodt (2015) "Testing Ambiguity Models through the Measurement of Probabilities for Gains and Losses," *American Economic Journal: Microeconomics* 7, 77–100.

Baillon, Aurélien, Zhenxing Huang, Asli Selim, & Peter P. Wakker (2018) "Measuring Ambiguity Attitudes for All (Natural) Events," *Econometrica* 86, 1839–1858.

Bruhin, Adrian, Helga Fehr-Duda, & Thomas Epper (2010) "Risk and Rationality: Uncovering Heterogeneity in Probability Distortion," *Econometrica* 78, 1375–1412.

Cappelli, Veronica Roberta, Simone Cerreia-Vioglio, Fabio Maccheroni, Massimo Marinacci, & Stefania Minardi (2018) "Sources of Uncertainty and Subjective Prices," working paper.

Cerreia-Vioglio, Simone, Fabio Maccheroni, & Massimo Marinacci (2019) "Ambiguity Aversion and Wealth Effects," *Journal of Economic Theory*, forthcoming.

Cerreia-Vioglio, Simone, Fabio Maccheroni, Massimo Marinacci, & Luigi Montrucchio (2011) "Uncertainty Averse Preferences," *Journal of Economic Theory* 146, 1275–1330.

Chateauneuf, Alain & José H. Faro (2009) "Ambiguity through Confidence Functions," *Journal of Mathematical Economics* 45, 535–558.

Chateauneuf, Alain, Jürgen Eichberger, & Simon Grant (2007) "Choice under Uncertainty with the Best and Worst in Mind: NEO-Additive Capacities," *Journal of Economic Theory* 137, 538–567.

Chew, Soo Hong, King King Li, Robin Chark, & Songfa Zhong (2008) "Source Preference and Ambiguity Aversion: Models and Evidence from Behavioral and Neuroimaging Experiments." *In* Daniel Houser & Kevin McGabe (eds.) *Neuroeconomics. Advances in Health Economics and Health Services Research* 20, 179–201, JAI Press, Bingley, UK.

Chierchia, Gabriele, Rosemarie Nagel, & Giorgio Coricelli (2018) "Betting 'on Nature' or 'Betting on Others': Anti-Coordination Induces Uniquely High Levels of Entropy," working paper.

Cozic, Mikael & Brian Hill (2015) "Representation Theorems and the Semantics of Decision-Theoretic Concepts," *Journal of Economic Methodology* 22, 292–311.

Dean, Mark & Pietro Ortoleva (2017) "Allais, Ellsberg, and Preferences for Hedging," *Theoretical Economics* 12, 377–424.

Dietrich, Franz (2018) "Savage's Theorem under Changing Awareness," *Journal of Economic Theory* 176, 1–54.

Dimmock, Stephen G., Roy Kouwenberg, Olivia S. Mitchell, & Kim Peijnenburg (2015) "Estimating Ambiguity Preferences and Perceptions in Multiple Prior Models: Evidence from the Field," *Journal of Risk and Uncertainty* 51 219–244.

Dimmock, Stephen G., Roy Kouwenberg, & Peter P. Wakker (2016) "Ambiguity Attitudes in a Large Representative Sample," *Management Science* 62, 1363–1380.

Dow, James & Sérgio R.C. Werlang (1992) "Uncertainty Aversion, Risk Aversion and the Optimal Choice of Portfolio," *Econometrica* 60, 197–204.

Eichberger, Jürgen & David Kelsey (2011) "Are the Treasures of Game Theory Ambiguous?," *Economic Theory* 48, 313–393.

Eichberger, Jürgen, Simon Grant, & Jean-Philippe Lefort (2012) "Generalized Neo-Additive Capacities and Updating," *International Journal of Economic Theory* 8, 237–257.

Ellsberg, Daniel (1961) "Risk, Ambiguity and the Savage Axioms," *Quarterly Journal of Economics* 75, 643–669.

Fehr-Duda, Helga & Thomas Epper (2012) "Probability and Risk: Foundations and Economic Implications of Probability-Dependent Risk Preferences," *Annual Review of Economics* 4, 567–593.

Ghirardato, Paolo, Fabio Maccheroni, & Massimo Marinacci (2004) "Differentiating Ambiguity and Ambiguity Attitude," *Journal of Economic Theory* 118, 133–173.

Ghirardato, Paolo & Massimo Marinacci (2001) "Risk, Ambiguity, and the Separation of Utility and Beliefs," *Mathematics of Operations Research* 26, 864–890.

Goeree, Jacob K. & Charles A. Holt (2001) "Ten Little Treasures of Game Theory and Ten Intuitive Contradictions," *American Economic Review* 91, 1402–1422.

Grant, Simon & Ben Polak (2013) "Mean-Dispersion Preferences and Constant Absolute Uncertainty Aversion," *Journal of Economic Theory* 148, 1361–1398.

Gul, Faruk & Wolfgang Pesendorfer (2014) "Expected Uncertain Utility Theory," *Econometrica* 82, 1–39.

Hansen, Lars P. & Thomas J. Sargent (2001) "Robust Control and Model Uncertainty," *American Economic Review, Papers and Proceedings* 91, 60–66.

Harvey, Charles M. (1990) "Structural Prescriptive Models of Risk Attitude," *Management Science* 36, 1479–1501.

Hey, John D., Gianna Lotito, & Anna Maffioletti (2010) "The Descriptive and Predictive Adequacy of Theories of Decision Making under Uncertainty/Ambiguity," *Journal of Risk and Uncertainty* 41, 81–111.

Izhakian, Yehuda (2017) "Expected Utility with Uncertain Probabilities Theory," *Journal of Mathematical Economics* 69, 91–103.

Kahneman, Daniel & Amos Tversky (1979) "Prospect Theory: An Analysis of Decision under Risk," *Econometrica* 47, 263–291.

Kilka, Michael & Martin Weber (2001) "What Determines the Shape of the Probability Weighting Function under Uncertainty," *Management Science* 47, 1712–1726.

Klibanoff, Peter, Massimo Marinacci, & Sujoy Mukerji (2005) "A Smooth Model of Decision Making under Ambiguity," *Econometrica* 73, 1849−1892.

Kocher, Martin G., Amrei Marie Lahno, & Stefan T. Trautmann (2018) "Ambiguity Aversion Is not Universal," *European Economic Review* 101, 268–283.

l'Haridon, Olivier, Ferdinand Vieider, Diego Aycinena, Augustinus Bandur, Alexis Belianin, Lubomir Cingl, Amit Kothiyal, & Peter Martinsson (2018) "Off the Charts: Massive Unexplained Heterogeneity in a Global Study of Ambiguity Attitudes," *Review of Economics and Statistics* 100, 664–677.

Lehrer, Ehud & Roee Teper (2015) "Subjective Independence and Concave Expected Utility," *Journal of Economic Theory* 158, 33–53.

Maccheroni, Fabio, Massimo Marinacci, & Aldo Rustichini (2006) "Ambiguity Aversion, Robustness, and the Variational Representation of Preferences," *Econometrica* 74, 1447–1498.

MacCrimmon, Kenneth R. & Stig Larsson (1979) "Utility Theory: Axioms versus "Paradoxes" ." *In* Maurice Allais & Ole Hagen (eds.) *Expected Utility Hypotheses and the Allais Paradox*, 333–409, Reidel, Dordrecht.

Machina, Mark J. & David Schmeidler (1992) "A More Robust Definition of Subjective Probability," *Econometrica* 60, 745–780.

Mackenzie, Andrew (2019) "A Foundation for Probabilistic Beliefs with or without Atoms," *Theoretical Economics* 14, 709–778.

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

Schmeidler, David (1989) "Subjective Probability and Expected Utility without Additivity," *Econometrica* 57, 571–587.

Stock, James H. & Mark W. Watson (2015) "*Introduction to Econometrics*." Pearson Education, Reading, Mass.

Trautmann, Stefan T. & Gijs van de Kuilen (2015) "Ambiguity Attitudes." *In* Gideon Keren & George Wu (eds.), *The Wiley Blackwell Handbook of Judgment and Decision Making* (Ch. 3), 89–116, Blackwell, Oxford, UK.

Tversky, Amos & Craig R. Fox (1995) "Weighing Risk and Uncertainty," *Psychological Review* 102, 269–283.

Tversky, Amos & Peter P. Wakker (1995) "Risk Attitudes and Decision Weights," *Econometrica* 63, 1255–1280.

Villegas, Cesáreo (1964) "On Quantitative Probability σ-Algebras," *Annals of Mathematical Statistics* 35, 1787–1796.

Viscusi, W. Kip & Wesley A. Magat (1992) "Bayesian Decisions with Ambiguous Belief Aversion," *Journal of Risk and Uncertainty* 5, 371–387.

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

Online Appendix of
“Belief Hedges: Applying Ambiguity Measurements to All Events and All Ambiguity Models”

Aurélien Baillon, Han Bleichrodt, Chen Li, & Peter P. Wakker
Erasmus School of Economics, Erasmus University Rotterdam, Rotterdam, the Netherlands, Baillon@ese.eur.nl, c.li@ese.eur.nl, Wakker@ese.eur.nl
February 2020

Online Appendix OA: Literature Surveys on Ambiguity

OA.1. Theoretical surveys of ambiguity

These include Etner & Tallon (2012), Gilboa & Marinacci (2016), Marinacci (2015), and Machina & Siniscalchi (2014). An empirical survey is Trautmann & van de Kuilen (2015).

OA.2. Papers with outcome-dependent ambiguity

These include Chew et al. (2008), Dobbs (1991), Gul & Pesendorfer (2014, 2015), He (2019), Kahneman & Tversky (1975), Klibanoff, Marinacci, & Mukerji (2005), Nau (2006), Neilson (2010), Olszewski (2007), Siniscalchi (2009), Skiadas (2013), and Skiadas (2015).

OA.3. Fields where ε-contamination model was used

These fields include decision theory (Baillon et al. (2018b; Basili, Chateauneuf, & Scianna 2018), robust statistics (Hodges & Lehmann 1952), finance (Epstein & Schneider 2010), insurance theory (Carlier, Dana, & Shahidi 2003), game theory (Aryal & Stauber 2014). It satisfies Chew and Sagi’s (2008) assumptions (Eq. 22), with a-neutral probabilities $P = Q$.

OA.4. ε in ε-contamination model taken as level of ambiguity

This happened in Alon & Gayer (2016), Chateauneuf, Eichberger, & Grant (2007), Gajdos et al. (2008), Ghirardato, Maccheroni, & Marinacci (2004; Proposition 6), Giraud (2014), Hill (2013), Klibanoff, Mukerji, & Seo (2014), Shattuck & Wagner (2016), and Walley (1991).

OA.5. Expected utility for risk & ambiguity neutrality equated with subjective expected utility

This happened in Chateauneuf & Faro (2009), Evren (2019), Ghirardato, Maccheroni, & Marinacci (2004), Ghirardato & Marinacci (2002), Klibanoff, Marinacci, & Mukerji (2005), Montesano & Giovannoni (1996), Nehring (1999), and Siniscalchi (2009).

OA.6. More ambiguity averse in Yaari-type way

Many papers 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). This implies identical risk attitudes and, once that is assumed, it suffices to take r above riskless. Papers include Chateauneuf and Faro (2009 p. 541), Dean and Ortoleva (2017 Definition 5), Epstein (1999 Eq. 2.3), Evren (2019), Frick,Iijima, and Le Yaouanq (2019), Ghirardato and Marinacci (2002 Definitions 4 & 7), Giraud (2014 Definition 7), Gul and Pesendorfer (2014 Corollary 1), Gul and Pesendorfer (2015 Propositions 3 and 4) with ideal events instead of risk, Klibanoff, Marinacci, and Mukerji (2005 Definition 5), Klibanoff, Mukerji, and Seo (2014 Definition 3.4), and Qu (2015). In particular, it is compatible with the pointwise ordering of the uncertainty aversion function G in Cerreia et al.'s (2011b Proposition 6) general uncertainty aversion model.

OA.7. Ambiguity indexes/premia in monetary units

This happened in Brenner & Izhakian (2018), Cubitt, van de Kuilen, & Mukerji (2018), Giammarino & Barrieu (2013), Izhakian & Brenner (2011), Jewitt & Mukerji (2017), Lang (2017), l'Haridon e al. (2018), Maccheroni, Marinacci, Ruffino (2013), and Montesano & Giovannoni (1996).

Online Appendix OB: Some Further Proofs

PROOF OF EQ. 10. The case of index b is clear. As for index a,

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

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

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

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

PROOF OF EQ. 12. 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 $\mathcal{H}$, $\frac{Cov(m,\nu)}{var(\nu)} = \sigma\frac{Cov(P,\nu)}{var(\nu)}$ = (by Eq. 25) σ. □

Online Appendix OC: Proofs for the Smooth Model (§4)

We analyze our indexes for the smooth ambiguity model. Our analysis is similar to Izhakian and Brenner (2011) who provided local ambiguity premiums expressed in monetary units. Our premiums are expressed in probability units. We fix $\gamma \succ \theta$ and analyze Eq. 1 under the smooth model of ambiguity, explaining notation later:

$$\int_{\Delta(S)} \varphi(Q(E))d\mu = \varphi(m(E)). \tag{OC.1}$$

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. OC.1, 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. OC.1 the first-order probability of receiving γ being certain to be $m(E). p = P(E) =$

$\int_{\Delta(S)} Q(E)d\mu$ denotes 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 ambiguity aversion (Klibanoff, Marinacci, and Mukerji 2005 p. 1865), and $o(\sigma^2)$ expresses first-order approximation as σ^2 vanishes.

LEMMA OC.1. For some given event E:

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

PROOF OF LEMMA OC.1. 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.

We assume in this Online Appendix OC that all functions are sufficiently smooth with all required derivatives existing and all O and o terms uniform. 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. OC.3 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, Marinacci, and Mukerji (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:

$$(F' \times (1 - \alpha, 1]) \ \cup\ (S' \times (0, (1 - \alpha)p]); \qquad \text{(OC.3)}$$

and θ otherwise. The events in Eq. OC.3 play the role of events E in Eq. OC.2 The limit of E tending to an ambiguity neutral event in the main text is achieved by letting α tend to 0 in Eq. OC.3. 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 $q = Q(F)$,

$$\int_{\Delta(S)} \varphi(\alpha q + (1-\alpha)p)d\mu = \varphi(m_\alpha)\,. \qquad \text{(OC.4)}$$

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

$$\varphi(m_\alpha) = \varphi(p) + (m_\alpha - p)\varphi'(p) + O((m_\alpha - p)^2) \qquad \text{(OC.5)}$$

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. OC.4 is:

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

$$\varphi(p) + \tfrac{1}{2}\alpha^2\varphi''(p)\tau^2 + o(\alpha^2) \qquad \text{(OC.6)}$$

(the term with φ' drops). Because of Eq. OC.4 we can equate Eqs. OC.5 and OC.6:

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

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

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

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

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

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

$\alpha^2\tau^2$ here is the variance of the event in Eq. OC.3 i.e., it is denoted σ^2 in Eq. OC.2 which now follows. □

Eq. OC.2 is Pratt's (1964) Eq. 5, but with probabilities replacing monetary outcomes. It illustrates once more that uncertainty about probabilities is treated in this analysis in the same way as uncertainty about outcomes was treated in traditional analyses.

Our aversion index b, which is Eq. OC.2 averaged over $\mathcal{H}$ and multiplied by 2 for normalization (see §2.2), we get Eq. 13:

$$b = \overline{\sigma^2 A(p)} + o(\sigma^2) \,. \tag{OC.7}$$

It is the product of what is sometimes interpreted as ambiguity perception (σ^2) and a relative index per perceived unit, $A(p)$. A similar decomposition occurred in Eq. 20, where it was discussed further. Eq. OC.7 makes the average of this product, involving the not directly observable p, A, and σ^2, directly observable because b is.

By Eq. OC.2, $a = 1 - \frac{Cov(p,v)}{Var(v)} + \frac{Cov\left(\frac{1}{2}\sigma^2 A(p) + o(\sigma^2), v\right)}{Var(v)}$. Because $Cov(p, v) = Var(v)$ (Eq. 25), we obtain Eq. 14. That is, the insensitivity index is:

$$a = \frac{1}{2}\frac{Cov(\sigma^2 A(p), v)}{Var(v)} + o(\sigma^2) \,.$$

It captures how the aversion premium (Eq. OC.2) increases with event size v, which indeed reflects insensitivity. This degree of ambiguity (perception), depending on variance of the event probability, is similar to Izhakian's (2017) measure in his variation of the smooth model that uses Choquet expected utility rather than expected utility in the second stage.

Online Appendix OD: Comparison of a with a′ in the experiment of Baillon et al. (2018)

The experiment of Baillon et al (2018) consisted of two treatments (control vs. time pressure) and two parts. Each scatter plot displays a' ($\frac{Cov(m,P)}{Var(P)}$) as a function of a ($\frac{Cov(m,v)}{Var(v)}$). Points that are not on the diagonal are all due to violations of the basic condition of monotonicity, so that our theoretical analyses do not apply to them.

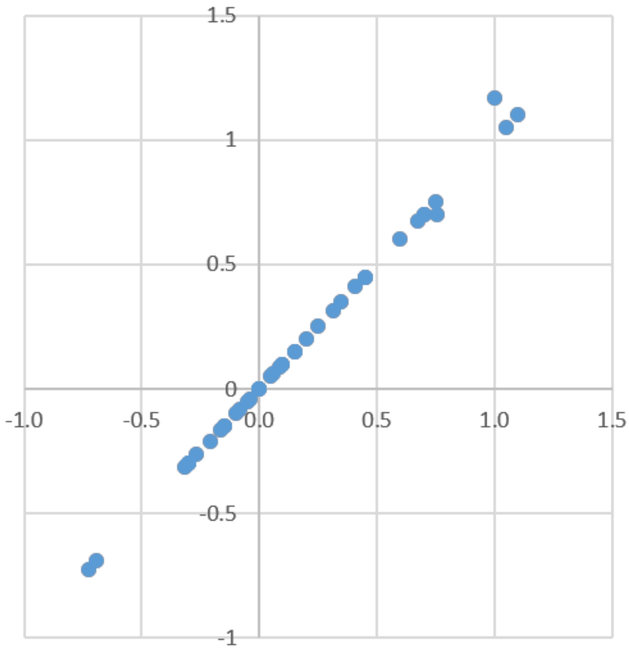

FIGURE O.1. a' as a function of a for the control treatment – Part 1

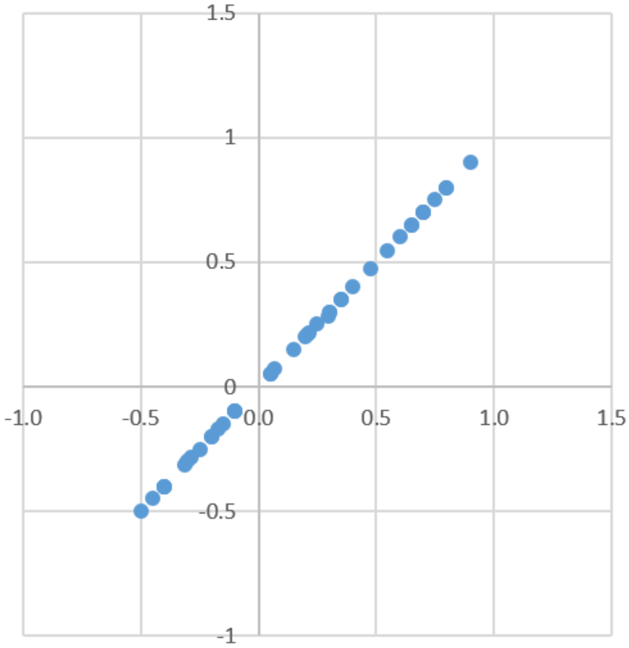

FIGURE O.2. a' as a function of a for the control treatment – Part 2

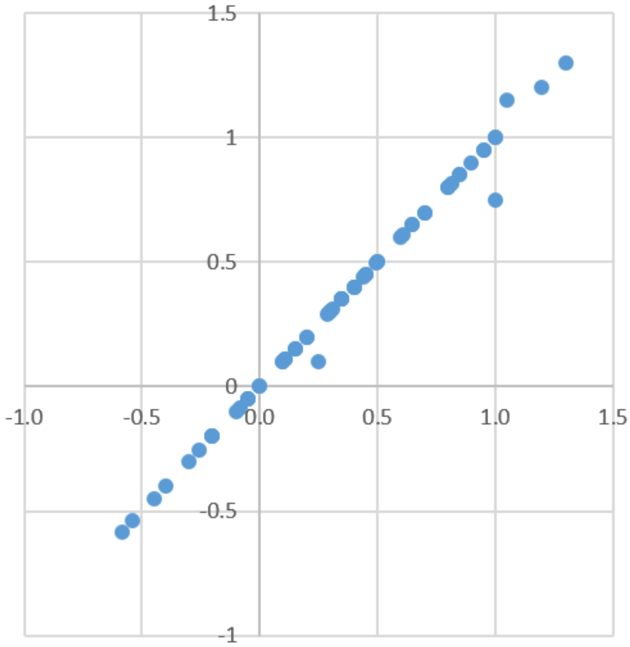

FIGURE O.3. Figure O.3. a' as a function of a for the time pressure treatment – Part 1

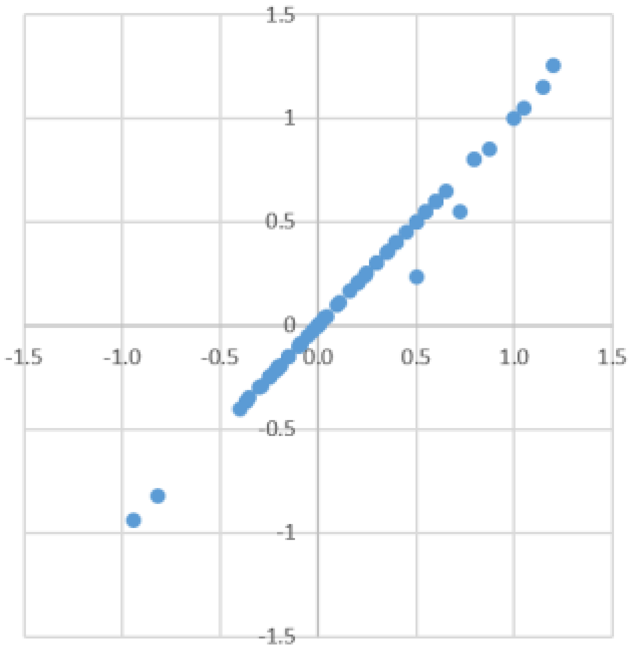

FIGURE O.4. Figure O.4. a' as a function of a for the time pressure treatment – Part 2

Online Appendix OE: Further References for Online Appendix

Alon, Shiri & Gabrielle Gayer (2016) "Utilitarian Preferences with Multiple Priors," *Econometrica* 84, 1181–1201.

Aryal, Gaurab & Ronald Stauber (2014) "Trembles in Extensive Games with Ambiguity Averse Players," *Economic Theory* 57, 1–40.

Baillon, Aurélien, Han Bleichrodt, Umut Keskin, Olivier L'Haridon, & Chen Li (2018b) "The Effect of Learning on Ambiguity Attitudes," *Management Science* 64, 2181–2198.

Basili, Marcello, Alain Chateauneuf, & Giuseppe Scianna (2018) "Coherent and Consistent Representation of Keynes's Long-Term Expectation," Working paper.

Brenner, Menachem & Yehuda Izhakian (2018) "Asset Pricing and Ambiguity: Empirical Evidence," *Journal of Financial Economics* 130, 503–531.

Carlier, Guillaume, Rose Anne Dana, & Niousha Shahidi (2003) "Efficient Insurance Contracts under Epsilon-Contaminated Utilities," *The Geneva Papers on Risk and Insurance Theory* 28, 59–71.

Cerreia-Vioglio, Simone, Paolo Ghirardato, Fabio Maccheroni, Massimo Marinacci & Marciano Siniscalchi (2011a) "Rational Preferences under Ambiguity," *Economic Theory* 48, 341–375.

Chateauneuf, Alain, Jürgen Eichberger, & Simon Grant (2007) "Choice under Uncertainty with the Best and Worst in Mind: NEO-Additive Capacities," *Journal of Economic Theory* 137, 538–567.

Chateauneuf, Alain & José H. Faro (2009) "Ambiguity through Confidence Functions," *Journal of Mathematical Economics* 45, 535–558.

Chew, Soo Hong, King King Li, Robin Chark, & Songfa Zhong (2008) "Source Preference and Ambiguity Aversion: Models and Evidence from Behavioral and Neuroimaging Experiments." *In* Daniel Houser & Kevin McGabe (eds.) *Neuroeconomics. Advances in Health Economics and Health Services Research* 20, 179–201, JAI Press, Bingley, UK.

Cubitt, Robin, Gijs van de Kuilen, & Sujoy Mukerji (2018) "The Strength of Sensitivity towards Ambiguity: A Qualitative Test and a Measurement," *Theory and Decision* 85, 275–302.

Dean, Mark & Pietro Ortoleva (2017) "Allais, Ellsberg, and Preferences for Hedging," *Theoretical Economics* 12, 377–424.

Dobbs, Ian M. (1991) "A Bayesian Approach to Decision-Making under Ambiguity," *Economica* 58, 417–440.

Epstein, Larry G. (1999) "A Definition of Uncertainty Aversion," *Review of Economic Studies* 66, 579–608.

Epstein, Larry G. & Martin Schneider (2010) "Ambiguity and Asset Markets," *Annual Review of Financial Economics* 2, 315–346.

Etner, Johanna, Meglena Jeleva, & Jean-Marc Tallon (2012) "Decision Theory under Ambiguity," *Journal of Economic Surveys* 26, 234–270.

Evren, Özgür (2019) "Recursive Non-Expected Utility: Connecting Ambiguity Attitudes to Risk Preferences and the Level of Ambiguity," *Games and Economic Behavior* 114, 285–307.

Frick, Mira, Ryota Iijima, & Yves Le Yaouanq (2019) "Boolean Representations of Preferences under Ambiguity," working paper.

Gajdos, Thibault, Takashi Hayashi, Jean-Marc Tallon, & Jean-Christophe Vergnaud (2008) "Attitude towards Imprecise Information," *Journal of Economic Theory* 140, 27–65.

Ghirardato, Paolo, Fabio Maccheroni, & Massimo Marinacci (2004) "Differentiating Ambiguity and Ambiguity Attitude," *Journal of Economic Theory* 118, 133–173.

Ghirardato, Paolo & Massimo Marinacci (2002) "Ambiguity Made Precise: A Comparative Foundation," *Journal of Economic Theory* 102, 251–289.

Giammarino, Flavia & Pauline Barrieu (2013) "Indifference Pricing with Uncertainty Averse Preferences," *Journal of Mathematical Economics* 49, 22–27.

Gilboa, Itzhak & Massimo Marinacci (2016) "Ambiguity and the Bayesian Paradigm." In Horacio Arló-Costa, Vincent F. Hendricks, & Johan F.A.K. van Benthem (eds.), *Readings in Formal Epistemology*, 385–439. Springer, Berlin.

Giraud, Raphael (2014) "Second order Beliefs Models of Choice under Imprecise Risk: Nonadditive Second Order Beliefs versus Nonlinear Second Order Utility," *Theoretical Economics* 9, 779–816.

Gul, Faruk & Wolfgang Pesendorfer (2014) "Expected Uncertain Utility Theory," *Econometrica* 82, 1–39.

Gul, Faruk & Wolfgang Pesendorfer (2015) "Hurwicz Expected Utility and Subjective Sources," *Journal of Economic Theory* 159, 465–488.

Harvey, Charles M. (1990) "Structural Prescriptive Models of Risk Attitude," *Management Science* 36, 1479–1501.

He, Ying (2019) "Revisiting Ellsberg and Machina's Paradoxes: A Two-Stage Evaluation Model under Ambiguity," working paper.

Hill, Brian (2013) "Confidence and Decision," *Games and Economic Behavior* 82, 675–692.

Hodges, Joseph L., Erich L. Lehmann (1952) "The Use of Previous Experience in Reaching Statistical Decisions," *Annals of Mathematical Statistics* 23, 396–407.

Izhakian, Yehuda & Menachem Brenner (2011) "The Uncertainty Premium in an Ambiguous Economy," *The Quarterly Journal of Finance* 1, 323–354.

Jewitt, Ian & Sujoy Mukerji (2017) "Ordering Ambiguous Acts," *Journal of Economic Theory* 171, 213–267.

Kahneman, Daniel & Amos Tversky (1975) "Value Theory: An Analysis of Choices under Risk," paper presented at the ISRACON conference on Public Economics, Jerusalem, 1975.

Klibanoff, Peter, Massimo Marinacci, & Sujoy Mukerji (2005) "A Smooth Model of Decision Making under Ambiguity," *Econometrica* 73, 1849–1892.

Klibanoff, Peter, Sujoy Mukerji, & Kyoungwon Seo (2014) "Perceived Ambiguity and Relevant Measures," *Econometrica* 82, 1945–1978.

l'Haridon, Olivier, Ferdinand Vieider, Diego Aycinena, Augustinus Bandur, Alexis Belianin, Lubomir Cingl, Amit Kothiyal, & Peter Martinsson (2018) "Off the Charts: Massive Unexplained Heterogeneity in a Global Study of Ambiguity Attitudes," *Review of Economics and Statistics* 100, 664–677.

Lang, Matthias (2017) "First-Order and Second-Order Ambiguity Aversion," *Management Science* 63, 1254–1269.

Maccheroni, Fabio, Massimo Marinacci, & Doriana Ruffino (2013) "Alpha as Ambiguity: Robust Mean-Variance Portfolio Analysis," *Econometrica* 81, 1075–1113.

Machina, Mark J. & Marciano Siniscalchi (2014) "Ambiguity and Ambiguity Aversion." *In* Mark J. Machina & W. Kip Viscusi (Eds.) "*Handbook of the Economics of Risk and Uncertainty*, Vol. 1," North-Holland, Amsterdam.

Marinacci, Massimo (2015) "Model Uncertainty," *Journal of the European Economic Association* 13, 1022–1100.

Montesano, Aldo & Francesco Giovannoni (1996) "Uncertainty Aversion and Aversion to Increasing Uncertainty," *Theory and Decision* 41, 133–148.

Nau, Robert F. (2006) "Uncertainty Aversion with Second-Order Utilities and Probabilities," *Management Science* 52, 136–145.

Nehring, Klaus D.O. (1999) "Capacities and Probabilistic Beliefs: A Precarious Coexistence," *Mathematical Social Sciences* 38, 197–213.

Neilson, William S. (2010) "A Simplified Axiomatic Approach to Ambiguity Aversion," *Journal of Risk and Uncertainty* 41, 113–124.

Olszewski, Wojciech (2007) "Preferences over Sets of Lotteries," *Review of Economic Studies* 74, 567–595.

Qu, Xiangyu (2015) "A Belief-Based Definition of Ambiguity Aversion," *Theory and Decision* 79, 15–30.

Shattuck, Mark & Carl Wagner (2016) "Peter Fishburn's Analysis of Ambiguity," *Theory and Decision* 81, 153–165.

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," *Theoretical Economics* 8, 59–93.

Skiadas, Costis (2015) "Dynamic Choice with Constant Source-Dependent Relative Risk Aversion," *Economic Theory* 60, 393–422.

Trautmann, Stefan T. & Gijs van de Kuilen (2015) "Ambiguity Attitudes." *In* Gideon Keren & George Wu (eds.), *The Wiley Blackwell Handbook of Judgment and Decision Making* (Ch. 3), 89–116, Blackwell, Oxford, UK.

Walley, Peter (1991) "*Statistical Reasoning with Imprecise Probabilities*." Chapman and Hall, London.