Randomization advice and ambiguity aversion

Christoph Kuzmics[1] · Brian W. Rogers[2] · Xiannong Zhang[3]

Abstract

We design and implement lab experiments to evaluate the normative appeal of behavior arising from models of ambiguity-averse preferences. We report two main empirical findings. First, we demonstrate that behavior reflects an incomplete understanding of the problem, providing evidence that subjects do not act on the basis of preferences alone. Second, additional clarification of the decision making environment pushes subjects' choices in the direction of ambiguity aversion models, regardless of whether or not the choices are also consistent with subjective expected utility, supporting the position that subjects find such behavior normatively appealing.

Keywords Knightian uncertainty · Subjective expected utility · Ambiguity aversion · Lab experiment

JEL Classification C91 · D81

1 Introduction

Many economic decisions are made under uncertainty that cannot be readily quantified by objective probabilities. Consider saving decisions, that is investing money into bonds or stocks in the presence of inflation and general uncertainties about the economic future. Even in the absence of wars and pandemics, most people find it hard to attach probabilities to the relevant possible events, let alone agree on such an assessment.

✉ Christoph Kuzmics
christoph.kuzmics@uni-graz.at

Brian W. Rogers
brogers@wustl.edu

Xiannong Zhang
zhangxiannong712@gmail.com

1 University of Graz, Graz, Austria

2 Washington University in St. Louis, 1 Brookings Dr., MO, St. Louis (63130), USA

3 Google LLC, 1600 Amphitheatre Parkway, Mountain View, CA 94043, USA

The classical paradigm of rational decision making under such uncertainty is subjective expected utility (SEU), underpinned by the axiomatic foundation of Savage (1954). The experimental designs of Ellsberg (1961), later implemented in many studies, see e.g., the survey of Trautmann and van de Kuilen (2015), have challenged the SEU paradigm. This challenge was mostly on positive, that is, empirical, grounds in that these experiments found that many people behave in a way that is inconsistent with subjective expected utility. The fact, however, that this behavior has been so robust in lab experiments and that many researchers have developed axiomatic foundations for alternative models of decision making that admit such behavior, see references below, may be received as posing a normative challenge to SEU in postulating that a broader set of preference models could be considered normatively appealing.

In this paper we aim to test the normative appeal of these alternative models of decision making under uncertainty. We use the term "normative" in the sense of Ellsberg's 1962 PhD thesis, see Ellsberg (2001, pages 22–26), and as in the "subjective" definition of rationality given by Gilboa (2012, page 5): We consider a decision normatively appealing (to the decision maker) if the decision maker (still) makes this choice after thorough reflection. In all of our treatments, subjects are provided with a complete, and fairly standard, description of all payoff-relevant aspects of the environment. We operationalize this reflection in the lab by providing subjects with supplementary descriptions that, while not payoff-relevant, emphasize certain ways to think about the environment. The descriptions are provided in the form of short videos that subjects watch before the elicitation of their choice. The experimental design has been pre-registered at the AEA RCT Registry, see Kuzmics et al. (2018).[1]

Our findings are relevant for all models of ambiguity aversion that are monotone. That is, if one act is better than another act in all states, then the inferior act cannot be chosen when the superior act is available. We call such models *classical ambiguity aversion (CAA) models*.[2]

The experimental environment is directly inspired by the hedging argument of Raiffa (1961) in the context of the Ellsberg (1961) two-color urn experiment. A *risky urn* contains 49 White and 51 Red balls.[3] An *ambiguous urn* also contains 100 balls, each of which is either Green or Yellow, but nothing more is known about the composition of the ambiguous urn. The decision-maker (DM) must choose one of three actions, after which the experimenter draws one ball from each urn. Together, these determine the consequence for the DM, which is either "Win" or "Lose", with

---

[1] The pre-registered experimental design comprises treatments for what are now two papers. Treatments 2 to 11 in Kuzmics et al. (2018) are reported here, while Treatment 1 is reported in Kuzmics et al. (2022).

[2] This category includes most preference models reviewed in the survey of Machina and Siniscalchi (2014), such as the maxmin expected utility model of Gilboa and Schmeidler (1989), the Choquet expected utility model of Schmeidler (1989), the smooth ambiguity model of Klibanoff et al. (2005), the variational and multiplier preference models of Maccheroni et al. (2006) and Hansen and Sargent (2001), confidence function preferences of Chateauneuf and Faro (2009), uncertainty aversion preferences of Cerreia-Vioglio et al. (2011), and the incomplete preference model of Bewley (2002).

[3] This variation from 50/50 avoids identification complications that can arise from indifference.

Win being strictly preferred to Lose. We call the choices bets for White, Green, and Yellow. Each bet Wins if the experimenter draws a ball of the corresponding color, and loses otherwise. Ellsberg's key insight is that Bet White, while commonly chosen, is incompatible with SEU.[4]

In this context, the Raiffa (1961) hedge against ambiguity works as follows. The DM flips a fair coin and bets Green if the coin lands on Heads and bets Yellow if the coin lands on Tails. This randomized action provides an (objective) winning probability of 50%, regardless of the color of the drawn ball from the ambiguous urn. As this is higher than the 49% winning probability from betting White, this action strictly dominates the Ellsberg choice.[5]

In light of the (Raiffa, 1961) argument, what are the possible explanations for a classically ambiguity-averse DM to nonetheless bet on White in this experiment even though it is dominated? First, it is possible, even likely, that designing a random choice does not occur to subjects as an option. Second, even if a subject recognizes such a possibility, it is possible that they cannot, or choose not to, go through the required construction and reasoning that would allow them to see that betting on White is dominated.[6] Suppose that a subject does recognize the possibility and understands the argument. A third possible explanation is that the subject has no access to a suitable randomization device, nor thinks they can simulate one. So suppose the subject does have a fair coin. The fourth and final explanation is that the subject lacks the ability to commit to the randomized action. Once the coin flip realizes, the subject could revisit their choice, and if they are ambiguity averse they will want to flip the coin again, and again ad infinitum, with one possible outcome that they bet on White in the end after all.

Classical ambiguity aversion models do not allow us to delve into the reasons behind the choice of White in the presence of the Raiffa (1961) argument, as these preference models are axiomatized for preferences over the space of all pure acts, see e.g., Seo (2009), and not over the set of all mixed acts as in Anscombe and Aumann (1963). This means, however, that whether or not a classically ambiguity averse subject who understands the environment may choose to bet on White depends simply on whether the Raiffa (1961) hedge (including the commitment to

---

[4] If Bet White is chosen, it must be weakly preferred to Bet Yellow, so that the DM's subjective probability of a Yellow ball being drawn from the ambiguous urn must be at most 49%. But then the subjective probability of a Green ball being drawn from the ambiguous urn must be at least 51%, so that Bet Green is strictly preferred to Bet White, a contradiction.

[5] Kuzmics (2017), appealing to results in classical statistical decision theory—in particular the complete class theorem of Wald (1947), has shown that a DM who can randomize over choices and commit to following the realized prescription of this randomization can never make choices that are inconsistent with SEU and at the same time consistent with CAA in any decision problem. See also Bade (2015), Oechssler and Roomets (2014), Azrieli et al. (2018), and Baillon et al. (2022a) for similar results and arguments along these lines. One can, in fact, regard ambiguity aversion as a preference for randomization, see Eichberger and Kelsey (1996) and Epstein and Schneider (2010).

[6] The argument is based on the reduction of compound lotteries and as Halevy (2007) found many subjects who make Ellsberg choices also fail to reduce compound lotteries. This failure, however, can hardly be seen as normatively appealing—it is simply a mathematical mistake. Abdellaoui et al. (2015) has found a much weaker association between displayed ambiguity aversion and a failure to reduce compound lotteries.

its outcome) is provided as a pure choice or not. If it is given as a pure choice the subject cannot choose to bet on White. If it is not given, they can.

Motivated by these considerations, our experimental treatments provide differential supplementary observations that focus on the hedging argument. All of the supplementary observations are given after a (common to all treatments) standard and complete description of the environment, and before the elicitation of the subject's choice. If the standard and complete description of the environment suffices to impart a full understanding of the consequences of each possible action, then the treatment effects of the supplementary observations will be null. Indeed, there is no marginal payoff-relevant information in the supplementary observations.

To the contrary, our first main finding is that the commentary significantly changes behavior. Since the commentary does not affect a subject's underlying preferences, it changes behavior through modifying a subjects' *understanding* of the environment. This is only possible if the subject's understanding after the standard description was incomplete or erroneous. We conclude that many subjects indeed have an incomplete or erroneous understanding after being provided with a standard and complete description of the classical Ellsberg two-urn environment, so that directly applying the revealed preference toolkit to such data may not be appropriate. Instead, a given choice may reflect an incomplete comprehension of the environment, and therefore be viewed as a mistake, or the result of confusion, rather than as a manifestation of the subject's preference.

Our second main finding concerns the direction of the effects of the supplemental observations. In this regard, and to the extent we can infer preferences from the treatments with commentary, our data supports the broader normative appeal of ambiguity aversion models over SEU in the following sense: When the Ellsberg choice (Bet White) is compatible with ambiguity averse preferences but not with SEU, the supplemental hedging observations increase the frequency of Ellsberg behavior; when the Ellsberg choice is incompatible with ambiguity aversion (and so also with SEU), the same observations decrease the frequency of Ellsberg behavior.

The paper proceeds as follows: The experimental design is given in Section 2. The results are given in Section 3. Section 4 offers a discussion of these results and Section 5 provides a brief survey of related literature before Section 6 concludes. Experimental instructions are presented in the Appendix.

2 Experimental design

The slight variation of the Ellsberg two-color urn experiment, outlined above, serves as the baseline and control. We then vary the supplemental observations that subjects receive about the decision-making environment. Our treatments differ from the baseline along two dimensions. First, in addition to the standard options of betting on White, Yellow, or Green, in some treatments subjects are presented with an additional fourth option, which executes a bet on either Green or Yellow according to the outcome of a fair coin to be tossed by a third party after the balls are drawn from the urns—the Raiffa hedge. Note that this option also serves as a commitment device for randomization, as the bet will be executed by the experimenter on the subject's

behalf. Recall that the compatibility of the Ellsberg choice (bet White) with classical ambiguity aversion models hinges on the presence or absence of this option.

Second, after the environment is fully described as transparently as possible, in some treatments subjects watch short videos before making their (single) choice. The videos, while all factually correct, emphasize different aspects of the consequences of using the randomization device, in ways that we hypothesize may change some subjects' understanding of the merits of the Raiffa hedge.

In the main treatment, the coin flip bet is included as an option, and subjects are presented with a single video (denoted V1) containing supplemental observations that describe the hedging argument of Raiffa (1961).[7] It describes the outcome of betting on the coin conditional on the outcome of the ball drawn from the ambiguous urn. It states that the winning probability using that option is 50% in either case (Green ball or Yellow ball drawn), and concludes by reminding the subject that betting on White wins with probability 49%.

This video, and each one of our other videos, does not explicitly advocate for any particular choice, so that it contains observations rather than advice. The transcripts of the videos are read by an anonymous (to subjects) third party to avoid a perception that the experimenters are giving implicit advice.

Partly to control for a possible experimenter demand effect, we designed a second video (denoted V2), in which the structure of the argument and the language is symmetric to the first video. It describes the outcome of betting on the coin conditional on the outcome of the coin flip. It states that no known winning probability can be specified in either case (Heads or Tails), and concludes by reminding the subject that betting on White wins with probability 49%. Again, it does not advocate for any particular choice.

We ran treatments utilizing exclusively this second video, as well as treatments in which subjects were presented with both videos before eliciting their choice (in both orders; there were no order effects).[8]

As we want to understand the effect of the supplementary observations independently from the effect of presenting the hedging device as an explicit option, we ran a parallel set of treatments with similar videos but where the available options were simply bets for White, Green, or Yellow, as in the baseline case, without the coin flip option. In these treatments, the Ellsberg choice remains compatible with CAA models. We varied the videos slightly to accommodate the different choice set. First, as there was no explicit coin, before showing either V1 or V2, we showed a preliminary video (denoted V0) in which it was explained that a subject could imagine a

---

[7] Transcripts of all videos are included in the Appendix.

[8] In sessions using both videos, we showed one video first, and the other video next. We did this in both orders. Then after both videos were shown, the subjects were provided with additional time to revisit any portions of either or both of the two videos before proceeding to enter their bet. During this time subjects could pause, rewind, and switch between videos as they liked. We thereby tried to minimize possible order effects and, indeed, there is no evidence that the order of the videos has any effect, and so we have pooled that data in our analysis.

Table 1 Summary of data (left: without a randomization device; right: explicit randomization option)

	W	G	Y		W	G	Y	C
Baseline	45%	41%	14%	Baseline	37%	26%	11%	26%
	(26/58)	(24/58)	(8/58)		(20/54)	(14/54)	(6/54)	(14/54)
V1	27%	55%	18%	V1	29%	2%	2%	67%
	(12/44)	(24/44)	(8/44)		(14/48)	(1/48)	(1/48)	(32/48)
V2	-	-	-	V2	41%	9%	9%	41%
					(18/44)	(4/44)	(4/44)	(18/44)
V1+V2	58%	28%	13%	V1+V2	23%	13%	9%	55%
	(35/60)	(17/60)	(8/60)		(13/56)	(7/56)	(5/56)	(31/56)

virtual coin toss, and then bet on Green/Yellow according to the outcome. Second, in videos V1 and V2 the coin toss was referred to as a virtual coin toss. We refer to the treatment with the explicit hedge/commitment option as "Coin" and those without it as "No Coin" treatments.

3 Results

Table 1 summarizes the main findings.

We summarize the key findings from this data in the following three results.

Result 1 If preferences alone dictate choices, the supplementary observations contained in the videos should have no effect on behavior. For the No Coin (Coin, resp.) treatments, pooling the data for Green and Yellow (being conservative), the p-value for the null hypothesis that choice frequencies are the same in the baseline and the $V1$ video treatment is 0.067 (< 0.001, resp.),[9] and for the null hypothesis that choice frequencies are the same in the baseline and the neutral $V1 + V2$ video treatment is 0.142 (0.007, resp.).

Result 2 The p-value for the null hypothesis that neutral video observations ($V1 + V2$) does not decrease the choice of White relative to the baseline is ≥ 0.5 for the No Coin treatments, and 0.057 for the Coin treatments.

Result 3 [10] Without supplemental observations there is no significant difference in the frequency of Bet White between the Coin and No Coin treatment, p-value 0.402. With supplemental observations ($V1 + V2$) there is a significant difference in the frequency of Bet White between the Coin and No Coin treatment, p-value < 0.001.

---

[9] All statistical tests performed in this paper are likelihood ratio tests.

[10] Readers should be aware that we initially planned only to test the two null hypotheses given in Results 1 and 2. The reported test provided here is an afterthought, after we have already seen the data, and is meant only to communicate that there is a significant (perhaps not meant in the statistical sense) and interesting difference of the effect of video commentary between the Coin and No Coin treatments.

4 Discussion

There is firm evidence that behavior across treatments is not a pure consequence of underlying preferences combined with a complete understanding of the environment. The observations contained in the videos cannot change preferences as they do not change any of the payoff-relevant considerations. Rather, any differences in behavior must come from differences in the subjects' understanding.

Suppose we adopt the view that choices after studying both videos indeed reveal preferences, since subjects may have a more complete understanding of the environment and the consequences of their choices after considering the observations contained therein. Even then, 23% of subjects, i.e., those who chose Bet White in the Coin treatment, have preferences different from any CAA model. The remaining 77% of subjects make choices in the Coin treatment that can be explained by CAA as well as SEU.

However, the No Coin treatments provide an interesting contrast. In these treatments, Bet White is undominated and the supplemental observations *increase* the frequency of Bet White relative to the baseline description. If, again, we view the choices after studying both videos as revealing preferences, the choice of Bet White made by 58% of subjects is inconsistent with SEU but is consistent with CAA models.

Together, these findings suggest that CAA models have broader normative appeal than (the narrower theory of) SEU despite their descriptive problems in some environments. Of note, we asked subjects (in a non-incentivized post-experiment questionnaire) if their preference became more or less clear after watching the videos. In the No Coin treatments 17% (27 out of 162) of subjects reported that their preferences became "less clear" after watching the videos, which is much higher than the 3% (6 out of 202) reporting the same in the Coin treatment ($p < 0.001$), calling into question the presumption that behavior directly reveals preferences, especially in the No Coin treatments. One interpretation is that many subjects find monotonicity to be a normatively appealing property, yet lack the sophistication to identify its consequences.

We conclude this discussion by considering preferences that may depend on the timing of the resolution of uncertainty, as in the models of (Seo, 2009; Saito, 2015; Ke & Zhang, 2020). Such models are not classical, as they are not monotone.[11] In the Coin treatments subjects were (truthfully) told, as part of the baseline description of the environment, that the coin flip would be executed after the balls were drawn from the urns and revealed. Thus, the choice of Bet White in the Coin treatments (37% in the baseline and 23% after both videos) is even inconsistent with these more flexible models.

Roughly, one could categorize our subjects as follows. There is one group of subjects ($\geq 23\%$) who make choices inconsistent with all models of ambiguity aversion. There is a second group of subjects ($\leq 58\% - 23\% = 35\%$) who make choices consistent with ambiguity aversion but not with SEU. The remainder make choices consistent with SEU. Subjects in the second group would have a demand for randomization devices as they cannot, to their satisfaction, create and commit to randomized choices themselves.

---

[11] Motivated by the state separability embedded in monotonicity, Bommier (2017) provides a model where monotonicity is relaxed.

5 Related literature

A number of papers have studied the consistency of subjects' choices across decision problems. These include Binmore et al. (2012), Stahl (2014), Voorhoeve et al. (2016), and Crockett et al. (2019). This literature finds, on the whole, that relatively few subjects make consistent choices, and those who do tend to be ambiguity-neutral. The lack of consistency can be interpreted as evidence against people choosing according to a clear preference. However, ambiguity-averse DMs may find inconsistent choices to be a useful hedge against ambiguity, see e.g., Kuzmics (2017) and Azrieli et al. (2018) for more general arguments.[12] Our single choice design is immune to such problems. This is why we constrained our design to a single incentivized elicitation per subject, even at the cost of forgoing the ability to conduct within-subject analyses across treatments.

Spears (2009), Dominiak and Schnedler (2011), and Oechssler et al. (2016) study experiments in which subjects are given the Raiffa hedge as an option, similarly to our baseline Coin treatment (without supplemental commentary). Generally, they find very few subjects choosing this option, with more subjects instead choosing a dominated option. This too is evidence against CAA models. They also find that subjects do not care about the timing of the resolution of uncertainty, evidence even against the non-CAA models of Seo (2009), Saito (2015), and Ke and Zhang (2020). Relatedly, Baillon et al. (2022b) study how subjects respond to two different forms of randomized incentive mechanisms to elicit ambiguity aversion preferences. Somewhat in contrast to the above mentioned literature, they find that, while about 50% of subjects make an Ellsberg-like choice in a treatment without randomization (akin to a treatment without a coin toss), only 25% to 29% of subjects make this choice when there is randomization (akin to a treatment with a coin toss). Interestingly, however, and consistent with the above mentioned literature, subjects do again not seem to care about the timing of the resolution of uncertainty.

We add to this literature by focusing on the possible effects on subjects' choices of providing explicit descriptions of the Raiffa hedge in the form of video commentaries. We thus add to these findings that such observations significantly influence behavior, and does so in directions that support the appeal of CAA models.

Nielsen and Rehbeck (2022) study the normative appeal of key axioms of rational decision-making under risk. They ask subjects two things: Whether they would agree that their choices should satisfy a certain axiom, and to make a choice from a set of lotteries. For all those subjects whose two answers are inconsistent, they point this out to them and ask them if they would reconsider their answers. They then find that of all those subjects whose answers are inconsistent, 47% revise their choice to be consistent with the axiom, while 13% revise their wish to be consistent with the axiom. Our approach has similarities with that of Nielsen and Rehbeck (2022). We also show subjects an axiom, if one can call it that, by giving them video

---

[12] This problem persists under many different preference elicitation schemes. See also Bade (2015), Oechssler and Roomets (2015), and Baillon et al. (2022a), which build on earlier work on eliciting non-expected utility preferences under objective uncertainty Karni and Safra (1987).

commentary about the consequences of the Raiffa (1961) hedge, i.e., of making choices based on a coin flip. We also have a similar goal as Nielsen and Rehbeck (2022): We are interested if subjects might make choices by mistake because they do not fully understand the consequences of their choices. In contrast to Nielsen and Rehbeck (2022), however, and this is the main difference, we here do not study decision-making under risk, but under ambiguity. Our "axiom" of interest, the Raiffa (1961) hedge, is, therefore, quite different. We think of the Raiffa (1961) hedge not so much as an axiom, but more of an argument. Therefore, we do not just ask subjects if they agree with the argument, but provide two neutrally phrased commentaries about the consequences of the coin toss, one implicitly advocating for and one against the Raiffa (1961) hedge. Some of our findings, while in a different domain of choice problems, are, nevertheless consistent with those of Nielsen and Rehbeck (2022): In both domains of decision-making under risk (Nielsen & Rehbeck, 2022) and decision-making under ambiguity (this paper) there is evidence of people making choices that they see as mistakes after induced contemplation. We then add to this finding evidence of the way these mistakes are made and in which way they are corrected for the domain of decision-making under ambiguity.

Finally, several studies test, in various other ways, the normative appeal of ambiguity aversion preference models.[13] The closest to our design is that of Slovic and Tversky (1974), who give subjects written advice for and against Allais (1953) and Ellsberg (1961) choices. However, their advice is built around the independence axiom, and so concerns a quite distinct domain. Jabarian and Lazarus (2022) also study the effect of a form of advice on subjects' decisions in a framework with ambiguity aversion. Their framework involves two independent draws from the same two-color ambiguous urn (and two draws from a 50-50 risky urn) in which many subjects make dominated choices, similarly also to Yang and Yao (2017) and Kuzmics et al. (2022). Subjects win if they draw two balls of the same color from the urn that they choose, making betting on the ambiguous urn a (weakly) dominant choice—as the more extreme the ball distribution in the ambiguous urn the higher the chance of drawing two balls of the same color. Jabarian and Lazarus (2022) have treatments in which subjects are given additional decision problems that should help them understand the mechanism why a choice is dominated. They find that while subjects do seem to understand the mechanism, they, nevertheless, do not seem to transfer this knowledge to the original problem in which they make dominated choices regardless. Finally, Keller et al. (2007), Trautmann et al. (2008), Charness et al. (2013), and Keck et al. (2014) study decision problems with ambiguity in groups (or under peer observation) and find, on the whole, that group discussion and related phenomena tend to lead to more ambiguity-neutral choices.

---

[13] Al-Najjar and Weinstein (2009) provide normative arguments against ambiguity aversion.

6 Conclusion

We have subjected classical preference models of ambiguity aversion models to tests of their normative appeal with experiments that stay close to the original Ellsberg (two-color urn) design.

We find that subjects' choices are affected by payoff-irrelevant commentary. This implies that at least one of the two treatments, without or with commentary, does not allow the full revelation of subjects' preferences.

At least some of our subjects do seem to see a certain normative appeal in the kind of behavior prescribed by classical models of ambiguity aversion and, in particular, the monotonicity axiom. Giving subjects access to additional commentary, in the form of short video clips, results in behavior that is more consistent with these models.

The nature of this normative appeal suggests that people, after sufficient reflection, would have a demand for the ability to commit to randomized choices, a demand which one would surmise should be observable. It would be interesting to identify instruments outside the lab, in the various areas of application of ambiguity aversion models, that could serve to satisfy this demand.

We also find that our subjects lack a complete and perfect understanding of their decision environment and how their choices map into final outcomes, in spite of the fact that we did our best to describe the environment completely and accurately. If this is the case, then there is room for further descriptions to influence behavior. We have shown that this is indeed readily observable, using the relatively weak instrument of short video clips providing commentary on the hedging argument of Raiffa. This means that in classical designs, it may be necessary to view a given choice as arising from a combination of preferences and how the subject understands the environment, where the second channel is non-trivial. Accordingly, a given choice may not provide direct evidence for or against any given preference model.

Appendix: Experimental design

Experiment details

The experimental sessions took place in April and May of 2018, and February of 2020. The experiment was conducted at the Experimental and Behavioral Economics Laboratory at University of California, Santa Barbara. There are two waves of data collection. In the first wave, 176 students participated in ten sessions and the average session length was 60 minutes. In the second wave, 213 students participated in 12 sessions and the average session length was 60 minutes. In all sessions, subjects answered exactly one incentivized question, which was related to guessing the color of a ball. If the guess was correct, the subject received 10 USD, and 0 USD otherwise. The show-up fee for all sessions was 5 USD. At the end of each session we conducted a short questionnaire. The questions were not incentivized, but we emphasized that answering these questions would be helpful for our research. The experiment was programmed using z-Tree (Fischbacher, 2007). See Fig. 2 for a screen shot.
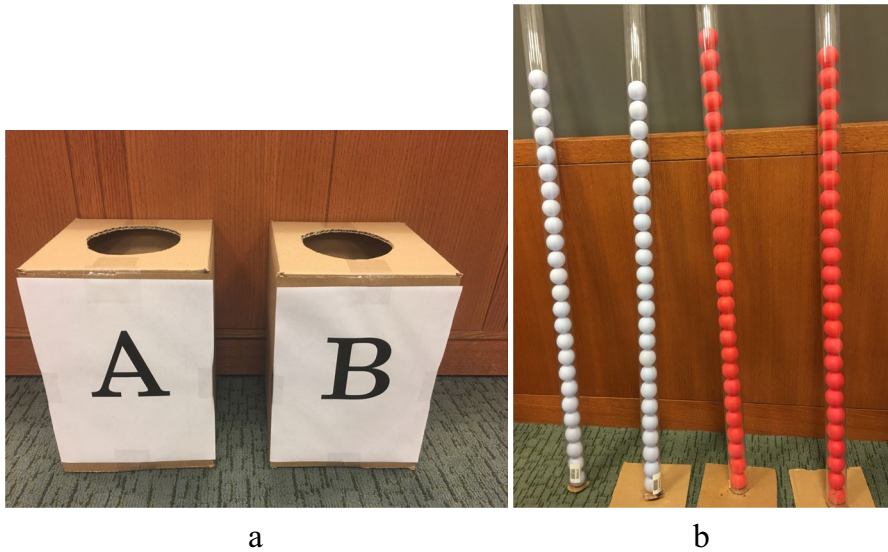
Fig. 1 Boxes

Physical environment

In all sessions, the urns and states were implemented using two cardboard boxes and colored ping-pong balls. During the experiment (and in what follows), we refer to the two containers as Box A and Box B. A photo of the boxes can be found in Fig. 1(a). The protocols we used were guided by the desire to be as clear and transparent as possible. Box A contained 49 White and 51 Red balls. The balls were displayed in clear plastic tubes at the beginning of the experiment so that subjects could easily see that there were two more Red than White balls. Photos of the tubes are included as Fig. 1(b). After showing the balls to subjects, they were poured into Box A. On the other hand, it was important that the exact contents of Box B were unknown. We therefore informed subjects that Box B contained 100 balls, each of which was either Green or Yellow, but we were intentionally not telling them anything further about the contents. Box B was shaken so that it was credible that it contained the same number of balls as Box A. After this presentation, subjects were told that they could inspect all the boxes and balls at the conclusion of the experiment if they so desired.

In each session, one subject was randomly selected to act as a monitor. The monitor was the person who physically conducted all draws of balls and displayed their colors to the other subjects, as well as coin flips, as relevant.
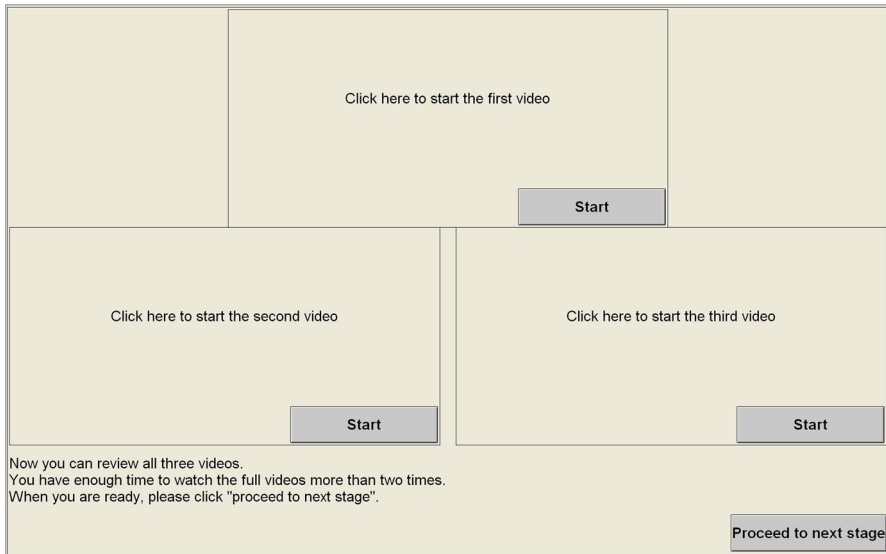
A screen shot

Fig. 2 Second experiment: Video review. After seeing videos and before making their choices, subjects had the chance to revisit all videos they watched before

Questionnaire

Table 2 shows the additional questions that we asked at the end of the experiment.

In the battery of sessions for the experiment, different treatments required different questions. The first block lists the questions that we asked in all treatments. The second block lists questions that are asked when both videos are presented. We denote by V1 the video in favor of the hedging argument and by V2 the video with the counter-argument. The third block lists questions that are asked when subjects are offered only three options and must implement the randomization with a virtual coin on their own. The fourth and fifth blocks list questions contingent on subjects' choices. In the treatments in which videos are not shown to the subjects before their incentivized choices, we showed the video before the questionnaire and asked for their "recommendation" in the questionnaire. We further asked for their reasoning. This is listed in the last block.

Instructions

We attached the instruction of the most comprehensive treatment. In this treatment, we provided the subjects with four options and both videos. Instructions for all the other treatments are written in a similar fashion.

Table 2 List of questions asked in the questionnaire

Questions asked in all groups

Gender (Male, Female, Prefer not to tell)

Major

How many Green balls do you think there are in Box B?

How many Yellow balls do you think there are in Box B?

Was there any part of the experiment that was unclear?

After watching the videos, my preference over choices was

(More clear, Less clear, Unchanged, I don't know)

Questions asked when V1 and V2 are both presented

Which video do you think is more compelling? (V1, V2, Equal)

Questions asked in when option Coin is not available

Do you find it difficult to simulate a coin toss in you head? (Yes, No)

Questions asked when subjects chose White ball in Box A

Why did you choose White?

Questions asked when subjects chose Green or Yellow ball in Box B

Why did you choose Green or Yellow?

Questions asked in no-video treatments

Why did you recommend this to your friend?

Instructions Welcome to the experiment! Please take a seat as directed. Please wait for instructions and do not touch the computer until you are instructed to do so. Please put away and silence all personal belongings, especially your phone. We need your full attention for the entire experiment. Adjust your chair so that you can see the screen in the front of the room. The experiment you will be participating in today is an experiment in decision making. At the end of the experiment you will be paid for your participation in cash. Each of you may earn different amounts. The amount you earn depends on your decisions and on chance. You will be using the computer for the experiment, and all decisions will be made through the computer. DO NOT socialize or talk during the experiment. All instructions and descriptions that you will be given in this experiment are factually accurate. According to the policy of this lab, at no point will we attempt to deceive you in any way. Your payment today will include a 5 dollar show up fee. One of you will be randomly selected to act as a monitor. The monitor will be paid a fixed amount for the experiment. The monitor will assist us in running the experiment and verifying the procedures. If you have any questions about the description of the experiment, raise your hand and your question will be answered out loud so everyone can hear. We will not answer any questions about how you "should" make your choices. As I said before, do not use the computer until you are asked to do so. When it is time to use the computer, please follow the instructions precisely.

We will now explain the experiment. There are two containers on the table that we will refer to as Box A and Box B. This is Box A. The Box is empty. Box A will contain 100 ping pong balls. Each of the balls in Box A will be either White, like this, or Red, like this. Specifically, Box A will contain exactly 49 White balls and 51 Red

balls, for a total of 100 balls. You don't have to remember these numbers. When it is time to make a decision, we will remind you of these numbers. We have counted and displayed the balls in these tubes to make it easier to show the contents of Box A. There are 25 White balls in this tube and 24 in this tube, for a total of 49 White balls. There are 25 Red balls in this tube and 26 Red balls in this tube, for a total of 51. We will now pour these balls into Box A and shake it to mix the balls together. This is Box B. We have already filled Box B with 100 ping pong balls. Each ball is either Green, like this, or Yellow, like this. We will not reveal the exact numbers of Green and Yellow balls. Instead, you know only that there are 100 balls in total, consisting of some combination of Green and Yellow balls. We will now shake Box B to mix the balls up. At the end of the experiment, you will have an opportunity to inspect the Boxes and ping pong balls if you wish. In a few moments we will ask the Monitor to draw one ball from each Box for everyone to observe. You will be asked to choose from several options that correspond to guessing the color of a ball that the Monitor draws. If your guess matches the result, you will receive 10 dollars in additional to the show up fee. If your guess does not match, you will receive 0 dollars in addition to the show up fee.

We will now start the experiment. On the computer desktop you will find a green icon named zleaf. Double click it now. Now there should be a welcome screen. Type your name and click the OK button in the welcome screen. One of you has been randomly selected by the software to serve as the monitor. Please raise your hand if you are the monitor. Could you please click the OK button on your screen and come to the front? Now your screen should have changed to "Please listen to the instructions". Please leave it like that and do not click OK. In a few moments the Monitor is going to draw one ball from Box A and one ball from Box B. We are going to ask you to bet on the outcome of those draws.

Specifically, you will be able to place one of four bets. Let me explain three of these options first. You can bet on White (from Box A), Green (from Box B) or Yellow (from Box B). Notice that you cannot bet on Red. If you bet on the White ball from Box A, then your payoff is not related to the draw from Box B. In other words, if the monitor draws a White ball from Box A, you win. If the monitor draws a Red ball, you lose. Similarly, if you choose Green or Yellow, your payoff only depends on the draw from Box B. For the fourth option, your bet will depend on the outcome of a coin flip. The monitor will flip a coin like this. If the coin lands on Heads, then we will set your bet to Green. If the coin lands on Tails, we will set your bet to Yellow. To repeat, we will set your bet to either Green or Yellow, depending on the coin flip result. Again, you don't have to write this down, since we will remind you about all the options when it is time to make your choice.

Before you make you decision, we are going to provide you with some comments on the experiment contained in two short videos. The videos are in total about 5 minutes long. After the videos, you will make your choice by selecting one of the four options. We will proceed like this: You will first watch the two videos. Then, you will have a chance to review the videos if you like. During the review session, you can pause or rewind the videos. There will be enough time to watch both videos more than two times in the review session. After the review session, we are going to ask for your choice. After you enter your decision, please wait for others to finish. There will not be any further instructions until all of you make your decisions.

Please follow the instructions on the screen and focus on the videos as much as possible. If you finish early, please remain quiet since others may still be watching. Now please put on the headphones provided at your desk and watch the videos. Once you are ready, please click OK.

The monitor is now going to draw the balls. Please look away and draw a ball from Box A and show it to everyone. The color is [REALIZED COLOR]. Please put the ball back. We will write down the result on the blackboard. Now please look away and draw a ball from Box B and show it to everyone. The color is [REALIZED COLOR]. Please put the ball back. We will write down the result on the blackboard. Please toss the coin and announce the result. The result is [REALIZED SIDE]. Please put the coin down. We will write down the result on the blackboard. Now please return to your seat and enter these results into your computer screen, accompanied by an Experimenter. You can now see the outcome and your earnings on the screen. If you have questions about your payoff, please raise your hand.

We will now conduct a short questionnaire. Please wait for the questionnaire to start. The monitor doesn't have to fill the questionnaire. Please complete the questionnaire. Please be as specific as you can in your responses. Answering the question is helpful to our research, but your responses are entirely voluntary. After you finish, please wait for others. We will call you to the front by your participant ID to be paid before leaving. Thank you very much for your participation. This concludes the experiment. We will now begin calling you to the front to be paid before leaving.

Video scripts

Names and notations

Recall that we denote the video that explains the Raiffa hedging argument, used in our main treatment, by V1 and its counter argument by V2. In the treatments without an explicit coin flip option, the instructions do not describe a coin. Instead, we show a short video introducing the hedging idea through the use of an "imaginary coin." We call this video V0. V0 is neither in favor of hedging nor against hedging. It merely states the idea of conditioning one's bet on the outcome of a virtual coin flip. We then slightly modified V1 and V2 by changing "the coin flip option" to "the imaginary coin." For more details, please see the script below.

To summarize, we have in total five distinct videos, listed in the table below, where "p" stands for "physical coin" and "v" for "virtual coin."

	V1p	V2p	V0	V1v	V2v
In favor of hedging?	Yes	No	n/a	Yes	No
Against hedging?	No	Yes	n/a	No	Yes
Describe hedging using a real coin?	Yes	Yes	n/a	No	No
Describe hedging using a virtual coin?	No	No	n/a	Yes	Yes
Description of a virtual coin?	n/a	n/a	Yes	n/a	n/a

V0 script

Recall that your three options are to choose: A White Ball from Box A, a Green Ball from Box B, or a Yellow Ball from Box B. Let me suggest a new method for choosing how to bet. To use this method, you need to create a random event. So, imagine you have a coin and you can flip it. The coin lands on Heads with probability 50% and on Tails with probability 50%. Before the toss, you plan to bet on a Green Ball from Box B if the coin lands on Heads, and on a Yellow Ball from Box B if the coin lands on Tails. Using this rule, you will not bet on a White Ball from Box A. To summarize, you first imagine the outcome of the coin flip. Then you choose Green Ball from Box B if the coin lands on Heads and you choose Yellow Ball from Box B if the coin lands on Tails.

V1p script

Recall that Box A contains 49 White balls and 51 Red balls. So, you will win with probability 49% if you choose the "White Ball from Box A." Let me describe the outcome when you choose the "Coin flip for Green/Yellow ball." Recall that Box B contains an unknown combination of 100 Green and Yellow balls. So when the Monitor draws a ball from Box B there are two possible cases: The ball can either be Green, or it can be Yellow. Suppose the ball happens to be Green. Now, when the monitor flips the coin, it will land either on Heads or on Tails. Each case is equally likely: The probability of Heads is 50% and the probability of Tails is 50%. If the coin lands on Heads, you would bet on Green and win. If the coin lands on Tails, you would bet on Yellow and lose. So, what we have observed is that if the ball from Box B happens to be Green, you would win with probability 50%. Now suppose that the ball from Box B happens to be Yellow. As before, when the monitor flips the coin, it will land either on Heads or on Tails. Each case is equally likely: The probability of Heads is 50% and the probability of Tails is 50%. If the coin lands on Heads, you would bet on Green and lose. If the coin lands on Tails, you would bet on Yellow and win. So, what we have observed now is that if the ball from Box B happens to be Yellow, you would again win with probability 50%. To summarize, if you choose the option "Coin flip for Green/Yellow ball" you will win with probability 50% whether the ball from Box B is Green or Yellow. Therefore, it does not matter how many of the balls are Green and how many are Yellow, since you will win with probability 50% in either case. By betting instead on a White Ball from Box A, you will win with probability 49%.

V2p script

Recall that Box A contains 49 White balls and 51 Red balls. So, you will win with probability 49% if you choose the "White Ball from Box A." Let me describe the outcome when you choose the "Coin flip for Green/Yellow ball." If you choose this option, there are two possibilities: When the monitor flips the coin, it will land either on Heads or on Tails. Each case is equally likely: The

probability of Heads is 50% and the probability of Tails is 50%. Suppose the coin happens to land on Heads. In this case, you would be betting on a Green Ball from Box B. The chance that betting on a Green Ball from Box B wins depends on how many Green balls are in Box B. Since you are not told how many Green balls are in Box B, your probability of winning is uncertain. So, what we have observed is that if the coin lands on Heads, your probability of winning is uncertain. Now suppose the coin happens to land on Tails. In this case, you would be betting on a Yellow Ball from Box B. The chance that betting on a Yellow Ball from Box B wins depends on how many Yellow balls are in Box B. Since you are not told how many Yellow balls are in Box B, your probability of winning is again uncertain.

V1v script

Recall that Box A contains 49 White balls and 51 Red balls. So, you will win with probability 49% if you choose the "White Ball from Box A." Let me describe the outcome when you choose the method based on the coin flip I described before. Recall that Box B contains an unknown combination of 100 Green and Yellow balls. So when the Monitor draws a ball from Box B there are two possible cases: The ball can either be Green, or it can be Yellow. Suppose the ball happens to be Green. Now, when you imagine flipping the coin, it will land either on Heads or on Tails. Each case is equally likely: The probability of Heads is 50% and the probability of Tails is 50%. If the coin lands on Heads, you would bet on Green and win. If the coin lands on Tails, you would bet on Yellow and lose. So, what we have observed is that if the ball from Box B happens to be Green, you would win with probability 50%. Now suppose that the ball from Box B happens to be Yellow. As before, when you imagine flipping the coin, it will land either on Heads or on Tails. Each case is equally likely: The probability of Heads is 50% and the probability of Tails is 50%. If the coin lands on Heads, you would bet on Green and lose. If the coin lands on Tails, you would bet on Yellow and win. So, what we have observed now is that if the ball from Box B happens to be Yellow, you would again win with probability 50%. To summarize, if you use the method based on the coin flip, you will win with probability 50% whether the ball from Box B is Green or Yellow. Therefore, it does not matter how many of the balls are Green and how many are Yellow, since you will win with probability 50% in either case. By betting instead on a White Ball from Box A, you have a known probability of winning of 49%.

V2v script

Recall that Box A contains 49 White balls and 51 Red balls. So, you will win with probability 49% if you choose the "White Ball from Box A." Let me describe the outcome when you choose the method based on the coin flip I described before. If you use this method, there are two possibilities: When you imagine flipping the coin, it will land either on Heads or on Tails. Each case is equally likely: The

probability of Heads is 50% and the probability of Tails is 50%. Suppose the coin happens to land on Heads. In this case, you would be betting on a Green Ball from Box B. The chance that betting on a Green Ball from Box B wins depends on how many Green balls are in Box B. Since you are not told how many Green balls are in Box B, your probability of winning is uncertain. So, what we have observed is that if your coin lands on Heads, your probability of winning is uncertain. Now suppose your coin happens to land on Tails. In this case, you would be betting on a Yellow Ball from Box B. The chance that betting on a Yellow Ball from Box B wins depends on how many Yellow balls are in Box B. Since you are not told how many Yellow balls are in Box B, your probability of winning is again uncertain. So, what we have now observed is that if your coin lands on Tails, your probability of winning is also uncertain. To summarize, if you use the method based on the coin flip, your probability of winning is uncertain if the coin lands on Heads and it is also uncertain if the coin lands on Tails. By betting instead on a White Ball from Box A, you have a known probability of winning of 49%.

Data availability All incentivized data is provided in Table 1. Full questionnaire data, as well as the videos, are available on request from the corresponding author.

References

Abdellaoui, M., Klibanoff, P., & Placido, L. (2015). Experiments on compound risk in relation to simple risk and to ambiguity. *Management Science, 61*(6), 1306–1322.

Allais, M. (1953). Le Comportement de l'Homme rationnel devant le Risque, Critique des Postulats et Axiomes de l'École Américaine. *Econometrica, 21*(4), 503–546.

Al-Najjar, N., & Weinstein, J. L. (2009). The ambiguity aversion literature: A critical assessment. *Economics and Philosophy, 25*(3), 249–284.

Anscombe, F. J., & Aumann, R. J. (1963). A definition of subjective probability. *Annals of Mathematical Statistics, 34*(1), 199–205.

Azrieli, Y., Chambers, C. P., & Healy, P. J. (2018). Incentives in experiments: A theoretical analysis. *Journal of Political Economy, 126*(4), 1472–1503.

Bade, S. (2015). Randomization devices and the elicitation of ambiguity-averse preferences. Journal of Economic Theory, 159(Part A), 221–235.

Baillon, A., Halevy, Y., & Li, C. (2022a). Experimental elicitation of ambiguity attitude using the random incentive system. *Experimental Economics, 25*(3), 1002–1023.

Baillon, A., Halevy, Y., & Li, C. (2022b). Randomize at your own risk: On the observability of ambiguity aversion. *Econometrica, 90*(3), 1085–1107.

Bewley, T. F. (2002). Knightian decision theory. Part I. *Decisions in Economics and Finance, 25*(2), 79–110.

Binmore, K., Stewart, L., & Voorhoeve, A. (2012). How much ambiguity aversion? *Journal of Risk and Uncertainty, 45*(3), 215–238.

Bommier, A. (2017). A dual approach to ambiguity aversion. Journal of Mathematical Economics, 71(Issue C), 104–118.

Cerreia-Vioglio, S., Maccheroni, F., Marinacci, M., & Montrucchio, L. (2011). Uncertainty averse preferences. *Journal of Economic Theory, 146*(4), 1275–1330.

Charness, G., Karni, E., & Levin, D. (2013). Ambiguity attitudes and social interactions: An experimental investigation. *Journal of Risk and Uncertainty, 46*(1), 1–25.

Chateauneuf, A., & Faro, J. H. (2009). Ambiguity through confidence functions. *Journal of Mathematical Economics, 45*(9–10), 535–558.

Crockett, S., Izhakian, Y. Y., & Jamison, J. C. (2019). *Ellsberg's hidden paradox. Baruch College Zicklin School of Business Research,* Paper No. 2019-07-04.

Dominiak, A., & Schnedler, W. (2011). Attitudes toward uncertainty and randomization: An experimental study. *Economic Theory, 48*(2–3), 289–312.

Eichberger, J., & Kelsey, D. (1996). Uncertainty aversion and preference for randomisation. *Journal of Economic Theory, 71*(1), 31–43.

Ellsberg, D. (1961). Risk, ambiguity, and the Savage axioms. *The Quarterly Journal of Economics, 75*(4), 643–669.

Ellsberg, D. (2001). *Risk, Ambiguity and Decision*. Garland Publishing.

Epstein, L. G., & Schneider, M. (2010). Ambiguity and asset markets. *Annual Review of Financial Economics, 2*(1), 315–346.

Fischbacher, U. (2007). z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics, 10*(2), 171–178.

Gilboa, I. (2012). *Rational choice*. MIT press.

Gilboa, I., & Schmeidler, D. (1989). Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics, 18*(2), 141–153.

Halevy, Y. (2007). Ellsberg revisited: An experimental study. *Econometrica, 75*(2), 503–536.

Hansen, L., & Sargent, T. J. (2001). Robust control and model uncertainty. *American Economic Review, 91*(2), 60–66.

Jabarian, B., Lazarus, S. (2022). *A two-ball Ellsberg paradox: An experiment*. https://doi.org/10.48550/arXiv.2206.04605

Karni, E., & Safra, Z. (1987). "Preference reversal'' and the observability of preference by experimental methods. *Econometrica, 55*(3), 675–685.

Keck, S., Diecidue, E., & Budescu, D. V. (2014). Group decisions under ambiguity: Convergence to neutrality. *Journal of Economic Behavior & Organization, 103*(Issue C), 60–71.

Keller, L. R., Sarin, R. K., & Sounderpandian, J. (2007). An examination of ambiguity aversion: Are two heads better than one? *Judgment and Decision Making, 2*(5), 390–397.

Ke, S., & Zhang, Q. (2020). Randomization and ambiguity aversion. *Econometrica, 88*(3), 1159–1195.

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

Kuzmics, C., Rogers, B. W., & Zhang, X. (2022). *An Ellsberg paradox for ambiguity aversion*. https://doi.org/10.48550/arXiv.2212.03603

Kuzmics, C. (2017). Abraham Wald's complete class theorem and Knightian uncertainty. *Games and Economic Behavior, 104*(Issue C), 666–673.

Kuzmics, C., Rogers, B. W., & Zhang, X. (2018). Choices in ambiguous settings. *AEA RCT Registry*. https://doi.org/10.1257/rct.2580-1.0.

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

Machina, M. J., & Siniscalchi, M. (2014). Ambiguity and ambiguity aversion. In M. Machina, W. K. Viscusi (Eds.), *Handbook of the Economics of Risk and Uncertainty* (pp. 729–807). Elsevier.

Nielsen, K., & Rehbeck, J. (2022). When choices are mistakes. *American Economic Review, 112*(7), 2237–2268.

Oechssler, J., Rau, H., & Roomets, A. (2016). *Hedging and ambiguity*. Discussion Paper Series: University of Heidelberg, Department of Economics.

Oechssler, J., & Roomets, A. (2014). Unintended hedging in ambiguity experiments. *Economics Letters, 122*(2), 243–246.

Oechssler, J., & Roomets, A. (2015). A test of mechanical ambiguity. *Journal of Economic Behavior & Organization, 119*(Issue C), 153–162.

Raiffa, H. (1961). Risk, ambiguity, and the Savage axioms: Comment. *The Quarterly Journal of Economics, 75*(4), 690–694.

Saito, K. (2015). Preferences for flexibility and randomization under uncertainty. *American Economic Review, 105*(3), 1246–1271.

Savage, L. (1954). *Foundations of Statistics*. John Wiley and sons.

Schmeidler, D. (1989). Subjective probability and expected utility without additivity. *Econometrica, 57*(3), 571–587.

Seo, K. (2009). Ambiguity and second-order belief. *Econometrica, 77*, 1575–1605.

Slovic, P., & Tversky, A. (1974). Who accepts Savage's axiom? *Behavioral Science, 19*(6), 368–373.

Spears, D. (2009). Preference for randomization? *Anscombe-Aumann inconsistency in the lab*. SSRN Working Paper.

Stahl, D. O. (2014). Heterogeneity of ambiguity preferences. *Review of Economics and Statistics, 96*(4), 609–617.

Trautmann, S. T., & van de Kuilen, G. (2015). Ambiguity attitudes. In G. Keren & G. Wu (Eds.), *The Wiley Blackwell Handbook of Judgment and Decision Making* (pp. 89–116). John Wiley & Sons Ltd.

Trautmann, S. T., Vieider, F. M., & Wakker, P. P. (2008). Causes of ambiguity aversion: Known versus unknown preferences. *Journal of Risk and Uncertainty, 36*(3), 225–243.

Voorhoeve, A., Binmore, K., Stefansson, A., & Stewart, L. (2016). Ambiguity attitudes, framing, and consistency. *Theory and Decision, 81*(3), 313–337.

Wald, A. (1947). An essentially complete class of admissible decision functions. *The Annals of Mathematical Statistics, 18*(4), 549–555.

Yang, C.-L., & Yao, L. (2017). Testing ambiguity theories with a mean-preserving design. *Quantitative Economics, 8*(1), 219–238.