



# Savage vs. Anscombe–Aumann: an experimental investigation of ambiguity frameworks

Jörg Oechssler<sup>1,3</sup> · Alex Roomets<sup>2</sup>

Accepted: 21 September 2020 / Published online: 9 October 2020  
© The Author(s) 2020

## Abstract

The Savage and the Anscombe–Aumann frameworks are the two most popular approaches used when modeling ambiguity. The former is more flexible, but the latter is often preferred for its simplicity. We conduct an experiment where subjects place bets on the joint outcome of an ambiguous urn and a fair coin. We document that more than a third of our subjects make choices that are incompatible with Anscombe–Aumann for any preferences, while the Savage framework is flexible enough to account for subjects' behaviors.

**Keywords** Ellsberg paradox · Ambiguity · Experiment

---

We would like to thank Adam Dominiak, Peter Dürsch, Jürgen Eichberger, Jean-Philippe Lefort, two anonymous referees, and participants of the conference on Ambiguity and Strategic Interactions in Grenoble in honor of Jürgen Eichberger for comments. Franklin and Marshall College provided financial support for the experiments.

---

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s11238-020-09778-w>) contains supplementary material, which is available to authorized users.

---

✉ Jörg Oechssler  
oechssler@uni-hd.de  
Alex Roomets  
roomets@gmail.com

<sup>1</sup> University of Heidelberg, Heidelberg, Germany

<sup>2</sup> Franklin and Marshall College, Lancaster, Pennsylvania, USA

<sup>3</sup> Department of Economics, University of Heidelberg, Bergheimer Str. 58, 69115 Heidelberg, Germany

## 1 Introduction

The Savage (1954) and the Anscombe and Aumann (1963) frameworks are the two most popular approaches when it comes to modeling ambiguity. The latter is a two-stage model where acts are maps from states to objective lotteries over consequences. It is often preferred for its simplicity, but the Savage model provides more flexibility. Gilboa and Schmeidler (1989) and Schmeidler (1989) used the Anscombe and Aumann approach as a basis for their seminal contributions to ambiguity theory. Eichberger and Kelsey (1996) show that, for standard ambiguity models like Choquet-expected utility (CEU) and Maxmin Expected Utility, ambiguity aversion implies a strict preference for randomization when looked at in the Anscombe–Aumann framework. They also show that the same need not hold in the Savage framework. Eichberger and Kelsey (1996) argue against the plausibility of a general preference for randomization but also admit the need for further experiments on this question.<sup>1</sup>

We implement an experiment in which some choices are inconsistent with ambiguity models that are based on the preference framework of Anscombe and Aumann (1963). We show that these choices can be consistent within a Savage framework using, e.g., a CEU model as in Eichberger and Kelsey (1996). The experiment involves subjects choosing from among six options that each relates to the outcomes of a coin flip and a draw from an ambiguous, 2-color urn. Two of the six options result in a clearly ambiguous act. Two more of the six options result in a clearly risky act. The last two options would be considered risky acts within the Anscombe–Aumann framework, but would be treated as ambiguous acts within the Savage framework. By manipulating the payoffs within the various acts, we are able to create a dominance relationship between the four risky acts using the Anscombe–Aumann framework. We find that dominated acts are still chosen by subjects more than a third of the time. The same subject choices can be explained with ambiguity models using the Savage framework, where the dominance relationship does not necessarily hold.

The two acts that highlight the differences between the two frameworks involve ambiguity hedging (see Oechssler and Roomets 2014, and Oechssler et al. 2019). These acts are akin to betting on one color when a coin flip comes up heads, and a different color when the coin flip comes up tails. Within the Anscombe–Aumann framework, subjects making such a combination exploit the complementarity of the probabilities of the two colors of balls in the urn to arrive at a believed 50:50 chance to win the bet. Within the Savage framework, such complementarity need not to be assumed. Subjects are allowed to believe that the probabilities of the two colors depend on the coin flip. Therefore, when a subject considers choosing an act that combines bets on blue (when the coins shows heads) and yellow (when the coin shows tails), the subject could believe that blue is unlikely when the coin shows heads and also that yellow is unlikely when the coin shows tails. Therefore, while

---

<sup>1</sup> In the meantime, a number of experiments (see in particular, Dominiak and Schmedler 2011 and Oechssler et al. 2019) have shown that, indeed, few subjects have a strict preferences for randomization.

the hedge acts represent risk using the Anscombe–Aumann framework, the same acts represent ambiguity using the Savage framework.

While it may seem we are pitting one framework against the other in a fair fight, we caution readers that the way we have been able to design choices leaves Savage mostly out of harms way while placing Anscombe and Aumann in jeopardy. Some may point out that the flexibility of the Savage framework is what keeps it out of the fray, and that this flexibility should be considered an advantage. We cannot disagree, but we leave discussions of the relative flexibility of the frameworks to more theoretical papers. As a fundamentally experimental endeavor, this paper should be viewed primarily as a test of the Anscombe–Aumann framework. Our results are not supportive of the Anscombe–Aumann framework in this context. This represents our main finding and contribution. It is, of course, interesting that the Savage framework could have explained our subjects' behavior when the Anscombe–Aumann framework could not. However, this should not be considered direct support for the Savage framework as there was no way it could have failed in our experimental setting.<sup>2</sup>

## 2 Experimental design

The experiment consisted of a single incentivized task,<sup>3</sup> followed by an unincentivized questionnaire. Subjects had to choose one of the six acts that depended on the outcome of a fair coin and the outcome of a draw from an Ellsberg urn.<sup>4</sup> The urn contained 24 blue and yellow balls in a composition that was unknown to subjects. Subjects were told that any combination from 0 blue balls (and 24 yellow balls) to 24 blue balls (and 0 yellow balls) was possible. Payoffs were chosen so as to ensure tie breaking for subjects who thought that some or all states are equally likely and to create the afore-mentioned dominance relationship within the Anscombe–Aumann framework. In treatment A, subjects chose from the six acts, as listed in Table 1. In treatment B, the payoffs of \$21 and \$22 were interchanged, while all other design aspects were kept constant. Interchanging the payoffs in this way helps us to identify the proportion of subjects choosing an option based on it having the highest potential payoff.

In the experiment, the acts were labeled neutrally “Option A” through “Option F” and were presented in a random order. Here, we have given them names that highlight their nature. The “heads” act, for example, will win if the coin shows heads, regardless of the ball draw. The “hedge yb” act would win if the ball drawn

---

<sup>2</sup> Here, one should also mention the intriguing thought experiments of Machina (2009, (2014) and the experiment of L'Haridon and Placido (2010), which make complementary but different points to our paper. These papers point out problems with CEU that are independent of whether the Savage or the Anscombe–Aumann framework is being used.

<sup>3</sup> Having several tasks with some probabilistic or fixed payment rule would run the risk of confounding ambiguity with hedging motives or with attitudes towards compound lotteries (see, e.g., Halevy 2007).

<sup>4</sup> In the actual experiment, we used a non-transparent bag and blue and yellow marbles. For expositional reasons, we employ the more customary urns and balls in the text.

**Table 1** Acts and payoffs

Acts	Coin shows heads		Coin shows tails	
	Ball blue $s_1$	Ball yellow $s_2$	Ball blue $s_3$	Ball yellow $s_4$
“Blue” ( $bb$ )	\$21	\$0	\$21	\$0
“Yellow” ( $yy$ )	\$0	\$21	\$0	\$21
“Heads” ( $h$ )	\$20	\$20	\$0	\$0
“Tails” ( $t$ )	\$0	\$0	\$20	\$20
“Hedge by” ( $by$ )	\$22	\$0	\$0	\$22
“Hedge yb” ( $yb$ )	\$0	\$22	\$22	\$0

is yellow and the coin shows “heads” or if the ball drawn is blue and the coin shows “tails”.

At the end of the experiment, subject volunteers drew a ball from the urn and tossed the fair coin. Importantly, the ball was drawn first (and shown to subjects), and then, the coin was tossed.<sup>5</sup> This timing was explained in the instructions.

After the acts were chosen, but before the random variables were determined, subjects filled out a questionnaire. The questionnaire included unincentivized questions about how subjects chose their bet in the elicitation task, a hypothetical three-color Ellsberg experiment, demographics, a hypothetical two-color Ellsberg urn, and beliefs about the random variables in the elicitation task (see the appendix for the questionnaire).

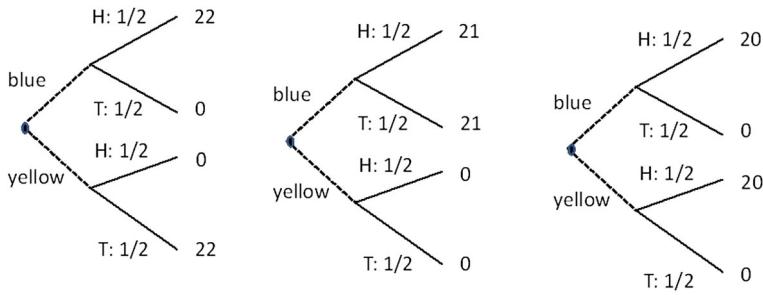
Experiments were conducted using pen and paper at the Economics Science Laboratory at the University of Arizona. Subjects were students at the university. There were 93 subjects in treatment A (57% female) and 31 subjects in treatment B (48% female). The experiment took roughly 30 min, and subjects received an average of \$19.91 including a \$10 show-up fee. Decisions and payments were made privately (with respect to other subjects).

Instructions (see Appendix) were distributed on paper and read aloud at the beginning of the experiment. Urns were on display during the entire experiment, so that subjects could be certain that the urns’ contents could not be manipulated. Subjects were allowed to verify the urns’ contents after the experiment, and some did.

### 3 Hypothesis

The two standard approaches to model uncertainty, the Anscombe and Aumann (1963) and the Savage (1954) framework, differ in the way they model a randomization device like a fair coin (see, e.g., Eichberger and Kelsey 1996, or

<sup>5</sup> This was done, so that we did not need to rely on the reversal-of-order axiom (see Anscombe and Aumann 1963). The timing of the coin toss theoretically has consequences (see, e.g., Seo 2009; or Eichberger et al. 2016), although it does not seem to matter experimentally (see Oechssler et al. 2019).



**Fig. 1** An illustration of a “Hedge” act (left), a “color” act (center), and a “coin” act (right)

Klibanoff 2001). In the Savage framework, the outcomes of a randomizing device must be modeled explicitly as part of the description of a state. The state space is the Cartesian product  $S = U \times R$ , where  $U = \{b, y\}$  is the outcome of the draw from an urn (ambiguous) and  $R = \{H, T\}$  is the outcome of a fair coin flip (objective randomization device). Hence, e.g.,  $s_1 = bH$  denotes the state where the drawn ball was blue and the coin flip produced heads. Thus, in our experiment, we have the state space  $S = \{s_1, \dots, s_4\}$ , as listed in Table 1, and a finite set of consequences  $X = \{0, 20, 21, 22\}$ . An act is a map  $f : S \rightarrow X$  and preferences are defined as binary relations on  $\mathcal{F}$ , the set of all acts.

In the experiment, there were the six acts, as listed in Table 1. Figure 1 illustrates the three types of acts available, the “hedge” acts ( $by, yb$ ), the “color” acts ( $bb, yy$ ), and the “coin” acts ( $h, t$ ). The tree to the left shows the “hedge by” act, the tree in center shows the act “blue”, and the tree to the right shows the act “heads”.<sup>6</sup>

In the Anscombe–Aumann framework, randomization devices are incorporated into the consequence space. The state space would consist only of  $S_{AA} = \{b, y\}$ . Consequences would be all simple lotteries (probability distributions) on  $X$ , denoted by  $\Delta(X)$ . Acts in the Anscombe–Aumann world are maps  $f : S_{AA} \rightarrow \Delta(X)$  and are listed in Table 2.

The crucial thing to note is that in an Anscombe and Aumann framework, both the “hedge” acts and the “coin” acts yield objective 50:50 lotteries. However, the hedge acts yield lotteries that pay out \$22 (\$21 in Treatment B) when successful, while the coin acts only pay out \$20. Thus, any decision-maker should strictly prefer either of the hedge acts to the coin acts.<sup>7</sup>

**Hypothesis** In the Anscombe–Aumann framework, no decision-maker should choose a coin act in either of the treatments.

<sup>6</sup> The remaining three acts are the mirror images of these three acts.

<sup>7</sup> Furthermore, in Treatment A, any ambiguity averse decision-maker with symmetric priors should strictly prefer the hedge acts to the color acts.

**Table 2** Acts and payoffs in Anscombe–Aumann

	Ball blue $s_1$	Ball yellow $s_2$
Hedge acts ( $by, yb$ )	$\frac{1}{2}\$22 + \frac{1}{2}\$0$	$\frac{1}{2}\$22 + \frac{1}{2}\$0$
Coin acts ( $h, t$ )	$\frac{1}{2}\$20 + \frac{1}{2}\$0$	$\frac{1}{2}\$20 + \frac{1}{2}\$0$
Color act $yy$	$\$0$	$\$21$
Color act $bb$	$\$21$	$\$0$

Note: In treatment B, the payoffs \$22 and \$21 are reversed

This hypothesis need not hold in a Savage framework (see Eichberger and Kelsey 1996). To construct a counter-example, consider a Choquet-Expected Utility (CEU) maximizer with the following capacity  $v(\cdot)$  and linear utility function  $u$ :

$$\begin{aligned}
 M \subseteq N &\Rightarrow v(M) \leq v(N) \\
 v(\emptyset) &= 0 \\
 v(S) &= 1.
 \end{aligned}$$

Following Eichberger and Kelsey (1996, Assumption 3.1), we assume that the capacity on  $S$  respects the probability of the coin flip for events that exclusively depend on the outcome of the coin flip. Under this assumption,  $v(\{s_1, s_2\}) = v(\{s_3, s_4\}) = 0.5$  and, therefore, coin acts are not ambiguous.

Now, suppose that  $v(\{s_i\}) = 0.1, \forall i, v(\{s_1, s_3\}) = v(\{s_1, s_4\}) = v(\{s_2, s_3\}) = v(\{s_2, s_4\}) = 0.2$  and  $v(\{s_1, s_2, s_3\}) = v(\{s_1, s_2, s_4\}) = v(\{s_2, s_3, s_4\}) = 0.6$ . In this case:

$$CEU(h) = 0.5u(20) = 10 > 0.2u(22) = 4.4 \geq CEU(f),$$

for all non-coin acts  $f$ .<sup>8</sup> Thus, a CEU maximizer need not satisfy the above hypothesis.

### 3.1 E-capacities

In fact, the above capacity is an example of the parametric capacity model of Eichberger and Kelsey (1999).<sup>9</sup> Their model offers a tractable way to incorporate an exogenous probability distribution into subjective (ambiguous) beliefs. For  $r \in R$ , let  $E_r := \{r\} \times U$  be the event referring to the outcome of the coin flip with known probability  $p(E_r) = \frac{1}{2}$  (i.e.,  $E_H = \{s_1, s_2\}$  and  $E_T = \{s_3, s_4\}$ ). An agent has information consistent probabilities  $\pi(s)$  if  $\sum_{s \in E_r} \pi(s) = p(E_r), r \in R$ . An example would be a uniform probability distribution  $\pi$  on  $S$ , such that  $\pi(s) = \frac{1}{4}$  for all  $s \in S$ .

The agent is confident that  $p(E_H) = p(E_T) = \frac{1}{2}$  describes the likelihoods of a fair coin. However, he is less confident about probabilities of states in  $E_H$  and  $E_T$ , respectively. That is, the agent distorts probabilities of states by his degree of confidence  $\rho_t \in [0, 1]$  and  $\rho_h \in [0, 1]$  that may vary across the known probability

<sup>8</sup> The inequality is due to the different payoffs in treatment A and B.

<sup>9</sup> We thank the editor, Adam Dominiak, for this observation.

events  $E_H$  and  $E_T$  (i.e.,  $\rho_t \neq \rho_h$ ). (Alternatively,  $\rho_t$  and  $\rho_h$  measure perceived ambiguity of states).

Formally, the EK capacity  $v^{EK}$  on  $2^S$  is defined as follows. For each  $A \in 2^S$ :

$$v^{EK}(A) := \sum_{r \in R} [\rho_r \pi(A \cap E_r) + (1 - \rho_r) p(E_r) \beta_r(A)],$$

$$\text{where } \beta_r(A) = \begin{cases} 1 & \text{if } E_r \cap A = E_r \\ 0 & \text{otherwise.} \end{cases}$$

When the degree of confidence is constant (i.e.,  $\rho_t = \rho_h$ ), the capacity is called the Ellsberg capacity. When  $\rho = 1$ , the EK capacity coincides with the probability measure  $\pi$  on  $S$ . Notice that the EK capacity is convex (hence, ambiguity aversion).

Consider a Choquet-Expected Utility preference with respect to an EK capacity, a strictly monotonic utility function  $u$ , and  $\pi(s) = \frac{1}{4}$  for all  $s \in S$ . Note that  $\pi(A \cap E_r) = \frac{1}{4}$  and  $\beta(A) = 0$ , for all  $A \in \{\{s_1, s_3\}, \{s_1, s_4\}, \{s_2, s_3\}, \{s_2, s_4\}\}$ . Then, the agent prefers a coin act to any non-coin act  $f$  as long as  $\rho_t + \rho_h < 2 \frac{u(20)}{u(22)}$ , since:

$$CEU(h) = CEU(t) = \frac{1}{2} u(20) > (\rho_t + \rho_h) \frac{1}{4} u(22) \geq CEU(f).$$

If  $\rho_t = \rho_h$ , then the preference for a coin ticket holds for any  $\rho < \frac{u(20)}{u(22)}$ .

### 4 Results

Subject decisions in our experiment are presented in Table 3. The left-hand side presents how many subjects chose the various acts, while the right-hand side combines acts of the same type and includes the percent of subjects choosing each type of act. The most important thing to notice is that there are many more coin act decisions than our main hypothesis would suggest. In fact, coin acts were the most popular choice when combining the data from both treatments. Statistically, this is a clear rejection of our main hypothesis. However, this hypothesis is very strict in that

**Table 3** Decision results by treatment

Decision	Treatment		Total	Type	Treatment		Total	
	A	B			A	B		
<i>bb</i>	11	10	21	Color acts	Count	24	15	39
<i>yy</i>	13	5	18		Percent	25.8%	48.4%	31.5%
<i>h</i>	19	5	24	Coin acts	Count	34	9	43
<i>t</i>	15	4	19		Percent	36.6%	29.0%	34.7%
<i>by</i>	19	3	22	Hedge acts	Count	35	7	42
<i>yb</i>	16	4	20		Percent	37.6%	22.6%	33.9%
Total	93	31	124	Total		93	31	124

a single coin act could be used to justify rejection. Therefore, it is worth considering whether coin acts could plausibly be explained as mistakes. If coin acts are a result of mistakes by subjects, otherwise, consistent with the Anscombe–Aumann framework, this would mean that (by a conservative estimate) around 1/3 of subjects made mistakes in our experiment. However, it would be more reasonable to assume that mistakes were randomly distributed over the choices subjects did not intend to make. Suppose that a share  $\lambda$  of subjects makes mistakes and deviates from their actually preferred act. If they make a mistake, they choose one of the remaining five acts with equal probability. Since a coin act should theoretically not be a preferred act, the two coin acts are always among the five remaining acts, yielding 2/5. Thus, to reproduce the share of coin acts of about one-third in the data, we need to have the share of mistakes  $\lambda$  solving  $1/3 = \lambda \frac{2}{5}$  or  $\lambda = 5/6$ . We believe that it is unlikely that 5/6 of our subjects made mistakes when indicating their preferred act, and so, we view our results as a strong rejection of our main hypothesis, even when allowing for some measurement error. Furthermore, results from the questionnaire, discussed further in Sect. 4.1, reject the notion that subjects were choosing at random.

#### 4.1 Who chose the coin acts?

While our main hypothesis and results concern the proportion of subjects that chose the various acts, we can also employ the questionnaire data to help explain why certain acts were chosen. For example, we look at what might have led subjects to choose a coin act, which is inconsistent with the Anscombe–Aumann framework. For each type of act, we estimate a linear probability model with a left-hand-side variable equal to “1” if the subject bet on that type of act and equal to “0” otherwise.<sup>10</sup> For explanatory variables, we use ambiguity attitude as measured separately by hypothetical two- and three-color Ellsberg urn questions in the questionnaire. We then use data from a written explanation of the original incentivized decision, which we asked for in the questionnaire.<sup>11</sup> To translate subjects’ written free-format explanations into a usable format, we employed three additional student coders who were asked to read through the questionnaire responses and identify whether certain topics were discussed. The topics included the relative “risk / safety”<sup>12</sup> and “known / unknown likelihood” of the different options, the idea that all options are equally likely, the relative payoffs of different options, and others.<sup>13</sup> These student coders entered a “1” if a topic was

<sup>10</sup> Logit and probit models yield similar conclusions.

<sup>11</sup> We asked subjects the following question immediately after choosing their incentivized bets and gave them a full page to respond: “What was your thought process when you made your decision?”

<sup>12</sup> This codes statements like, “option A seemed riskier than option B” or, “the safest choice was option D”.

<sup>13</sup> A full list of topics and the instructions given to the student coders is available as an appendix. Coders had access to the experimenters while working to ask clarifying questions about the topics, but the experimenters declined to answer questions about how to code specific responses.



discussed, and a “0” otherwise. The three codings were averaged to create our final measure of topics discussed, which we use in our regressions.<sup>14</sup>

Since payoffs of the various acts differ across treatments, we also use a treatment dummy and a treatment dummy interacted with a questionnaire response related to payoff comparisons. Subjects who were choosing based on payoff maximization potential would have chosen a hedge act in treatment A, but a color act in treatment B. The interaction term, therefore, allows for a reasonable interpretation of the “relative payoff” coding across treatments, particularly when looking at the color or hedge columns.<sup>15</sup> Finally, we also include a gender dummy, since the gender composition of the two treatments varied slightly.

Regression results are available in Table 4. We find that choosing a coin act seemed to be preferred by subjects that expressed ambiguity averse preferences in the hypothetical 3-color urn question,<sup>16</sup> cited the relative risk and/or the relatively known likelihoods of various outcomes, and did not comment on all outcomes being equally likely. Relative payoff discussion is negatively correlated to choosing coin acts (which had the lowest expected payoff) but not significantly so.

As one might suspect, choosing a color act was negatively correlated to ambiguity aversion measures (both three-color and two-color measures).<sup>17</sup> Choosing a color act was also negatively correlated with the discussion of payoff differences in treatment A, when “Hedge betting” had the highest winning payoff. In treatment B, when the color acts had the highest winning payoff, this effect is cancelled out.

Choosing a hedge act seemed to be preferred by subjects that discussed that acts were equally likely, and focused on payoff differences. This led to much more frequent hedge act choices in treatment A, where the hedge acts had the highest winning payoff. Unlike what would be expected according to theory, choosing hedge acts did not appear particularly related to ambiguity aversion.

It is also worth pointing out that the systematic and intuitively consistent differences highlighted in this section provide evidence that subjects both understood the options they were given in the main experiment and made their choices based on intuitively reasonable criteria. Indeed, it is unlikely that subjects made their decision randomly due to either misunderstanding or indifference.

---

<sup>14</sup> Codings of the three coders were unanimous about 92% of the time.

<sup>15</sup> The coefficient for “Relative Payoffs” can be interpreted as the effect in treatment A, while the sum of the coefficients for “Relative Payoffs” and “Relative Payoff  $\times$  Treatment” can be interpreted as the effect in treatment B.

<sup>16</sup> Preferences with respect to the 2-color urn were positively correlated with coin acts, but this relationship was not significant.

<sup>17</sup> The two-color measure is only marginally significant. This may be because there is always the possibility of misclassifying ambiguity averse subjects if they have asymmetric priors.

**Table 4** Linear probability regressions of questionnaire responses

Variable	Probability of act type		
	Color	Coin	Hedge
Ambiguity averse (3 color)	-0.194** (0.082)	0.185** (0.079)	0.009 (0.074)
Ambiguity averse (2 color)	-0.212 <sup>†</sup> (0.116)	0.165 (0.112)	0.047 (0.105)
Risk / safety	-0.122 (0.130)	0.399*** (0.125)	-0.276** (0.117)
Known / unknown	-0.129 (0.104)	0.434*** (0.100)	-0.305*** (0.094)
Equally likely	-0.085 (0.130)	-0.435*** (0.125)	0.520*** (0.117)
Relative payoffs	-0.227** (0.105)	-0.125 (0.101)	0.353*** (0.095)
Treatment (treatment B = 1)	0.135 (0.111)	-0.262** (0.107)	0.128 (0.100)
Relative payoff × treatment	0.288 (0.225)	0.141 (0.217)	-0.429** (0.203)
Gender (female = 1)	0.137 <sup>†</sup> (0.078)	-0.098 (0.075)	-0.039 (0.071)
Constant	0.556*** (0.193)	0.232 (0.186)	0.213 (0.174)
<i>N</i>	124	124	124
Adjusted <i>R</i> <sup>2</sup>	0.170	0.267	0.349

\*, \*\*, \*\*\* Significant at the 10%, 5%, and 1% level, respectively. Standard errors in parentheses

## 5 Conclusion

Based on our hypothesis, subjects in our experiment should not have chosen coin acts according to the Anscombe–Aumann framework. However, more than one-third of our subjects did. Given that coin acts made up precisely one-third of the options available to subjects, attributing these choices to measurement error would imply that practically all subjects erred in their selection or were indifferent between options (despite the payoff asymmetry). The latter seems particularly unlikely given the results from the questionnaire that evidence a sensible pattern of preferences; Ambiguity averse subjects chose the coin acts more often than ambiguity neutral/loving subjects. Therefore, we are left to assume that subjects expressed a meaningful preference for the coin acts, contradicting our hypothesis. Many subjects, it seems, did not view both the coin and hedge acts as 50/50 propositions, or, if they did, there was some other factor that affected preferences, but was not

modeled. Either way, models using the Anscombe–Aumann framework were unable to correctly explain a large portion of subject decisions in our setting.

While it may seem then that we are left endorsing the Savage framework, we stop short of such an endorsement. While we do not find violations within the choice data, since any choice is plausibly supported in the Savage framework, neither do we find positive support for the framework in subjects' written explanations of their act choices. When asked to explain the act they chose, fewer than 5% of subjects were coded as having discussed the combined outcomes that make up the state space in the Savage framework. Instead, subjects tended to reference the coin flip and the ball draw independently. Of course, subjects need not express the particulars of a framework in writing to employ that framework in their decision-making. Therefore, we see our results as neutral with respect to the Savage framework, and leave the door open to the possibility that subjects adhere to a framework which we failed to consider in this paper.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

**Funding** Open Access funding enabled and organized by Projekt DEAL.

## References

- Anscombe, F. J., & Aumann, R. J. (1963). A Definition of Subjective Probability. *Ann. Math. Stat.*, *34*, 199–205.
- Dominiak, A., & Schmedler, W. (2011). Uncertainty aversion and Preferences for Randomization: An Experimental Study. *Econom Theory*, *48*(2), 289–312.
- Eichberger, J., & Kelsey, D. (1996). Uncertainty aversion and preference for randomisation. *J. Econ. Theory*, *71*(1), 31–43.
- Eichberger, J., & Kelsey, D. (1999). E-capacities and the Ellsberg paradox. *Theory Dec.*, *46*(2), 107–138.
- Eichberger, J., Grant, S., & Kelsey, D. (2016). Randomization and dynamic consistency. *Econ. Theory.*, *62*, 547–566.
- Ellsberg, D. (1961). Risk, ambiguity, and the Savage axioms. *Quart. J. Econ.*, *75*, 643–669.
- Gilboa, I., & Schmedler, D. (1989). Maxmin expected utility with non-unique prior. *J. Mathem. Econ.*, *18*, 141–153.
- Halevy, Y. (2007). Ellsberg Revisited: An Experimental Study. *Econometrica*, *75*, 503–536.
- Klibanoff, P. (2001). Stochastically Independent Randomization and Uncertainty Aversion. *Econom. Theory.*, *18*(3), 605–620.
- L'Haron, O., & Placido, L. (2010). Betting on Machina's reflection example: An experiment on ambiguity. *Theory. Dec.*, *69*(3), 375–393.
- Machina, M. J. (2009). Risk, ambiguity, and the rank-dependence axioms. *Am. Econ. Rev.*, *99*(1), 385–92.
- Machina, M. J. (2014). Ambiguity aversion with three or more outcomes. *Am. Econ. Rev.*, *104*(12), 3814–40.
- Oechssler, J., Rau, H., & Roomets, A. (2019). Hedging, ambiguity, and the reversal of order axiom. *Games Econ. Behav.*, *117*, 380–387.

- Oechssler, J., & Roomets, A. (2014). Unintended Hedging in Ambiguity Experiments. *Econ. Lett.*, *122*, 243–246.
- Savage, L. (1954). *Found. Stat.* New York: Wiley.
- Schmeidler, D. (1989). Subjective probability and expected utility without additivity. *Econometrica*, *57*, 571–587.
- Seo, K. (2009). Ambiguity and Second-order Belief. *Econometrica*, *77*(5), 1575–1605.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.