AMBIGUITY AVERSION IN CAPUCHIN MONKEYS

Jing Pan

(Under Direction of Dorothy Fragaszy)

Abstract:

An ambiguous event happens at multiple probability levels or within a range of probability levels; a risky event happens at a single probability level. Expected utility theory, subjective expected utility theory, and prospect theory from human economics predict no difference in choice of risky vs ambiguous events. Stimulus response models applied to choice behavior of nonhuman animals predict risk aversion. Contrary to these predictions, three tufted capuchin monkeys (*Sapajus apella*) were averse to ambiguous events over risky events. Our results are partially explained by our novel behavioral economic model, Pan-Fragaszy Softmax model. Our results are the first to show ambiguity aversion in monkeys, and because they are contrary to the prediction of the stimulus-response model, they strengthen the evidence for meta-cognition in nonhuman animals.

INDEX WORDS: uncertainty, decision making, meta-cognition, behavioral economics, risk, ambiguity.
AMBIGUITY AVERSION IN CAPUCHIN MONKEYS

by

Jing Pan

B.S. University of Science and Technology of China, China, 2005

M.S. University of Georgia, USA, 2009

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in
Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2013
ACKNOWLEDGEMENTS

I am grateful that the Psychology Department at UGA and UGA Graduate School Dean’s Award provides funding for this study. The IACUS approval number of this project is AUP #: A2012 02-026-Y1-A0. I appreciate help received from the animal care facility at UGA life science building and UGA instrument shop. I personally receive help on programming from Mr. Mingjing Zhang, Mr. Ebenezer Abboah-afari, Dr. Nan Li, and Dr. Yan Wang, on designing electronic devices from Dr. Lei Wang, and mathematics from Dr. Cheolwoo Park and Dr. Chuan Zeng. I thank my committee member for academic help along the years. At last, I thank UGA karate club and UGA Health Center for keeping me alive.
Table of contents

Pages

ACKNOWLEDGEMENT ...................................................................................................................................... iv

CHAPTER 1 ........................................................................................................................................................ 1

Introduction and literature review ..................................................................................................................... 1

   Literature review ........................................................................................................................................... 1

   Introduction .............................................................................................................................................. 4

Method .............................................................................................................................................................. 7

Results ............................................................................................................................................................. 9

Discussion ....................................................................................................................................................... 10

Conclusions ..................................................................................................................................................... 19

Appendices .................................................................................................................................................... 20

   Appendix 1 Stimulus response model cannot predict ambiguity aversion .............................................. 20

   Appendix 2 Appendix Method .................................................................................................................. 22

   Appendix 3 Appendix results .................................................................................................................. 35

   Appendix 4 Pan-Fragaszy softmax model explains ambiguity aversion in the decline paradigm and urn paradigm ................................................................................................................................. 42

   Appendix 5 Challenges of Kahn and Sarin model .................................................................................. 45

References ...................................................................................................................................................... 49
CHAPTER 1

Introduction and literature review

Literature review

Ambiguity aversion and how it has been evaluated.

Uncertainty is the probability of an uncertain event (Tversky & Kahneman, 1982). There are two types of uncertainty: risk and ambiguity (Hsu, Bhatt, Adolphs, Tranel, & Camerer, 2005). Economists define ambiguity as any distribution of probabilities other than a point estimate of probability (risk) (Becker & Brownson, 1964). While the degree of uncertainty is measured by the best estimated subjective probability (i.e., the mean judged probability if the distribution of the judged probabilities is symmetric), the degree of ambiguity is measured by the spread or range of the judged probabilities (Becker & Brownson, 1964): the wider the range, the more ambiguous the uncertainty.

When the degree of uncertainty is the same, healthy human subjects prefer risky uncertainty instead of ambiguous uncertainty, an empirical phenomenon called ambiguity aversion. The ambiguity aversion is best illustrated in Ellsberg’s paradox (Ellsberg, 1961). A simplified version of Ellsberg paradox is illustrated as the follows. Imagine one deck of 20 cards composed of 10 red and 10 blue cards (the risky deck). Another deck has 20 red or blue cards, but the composition of red and blue cards is completely unknown (the ambiguous deck). A bet on a color of either red or blue pays a fixed sum (e.g., $10) if a card with the chosen color is drawn and zero otherwise. There are two slightly different experimental paradigms that compare people’s behaviour when they face betting on a risky deck or an ambiguous deck. In the first
paradigm, when subjects can get $3 if they drop the bet and in each bet the component of the ambiguous deck is new but unknown, healthy human adults proportionally drop more often when they bet from the ambiguous deck than from the risky deck (Hsu et al., 2005). In the second paradigm, subjects do not have a fixed payoff for dropping a trial, but they were asked to choose which deck they are more willing to bet from. Healthy human adults are more willing to bet from the risky deck compared to the ambiguous deck (Becker & Brownson, 1964; Loewenstein, Rick, & Cohan, 2008). The results in both paradigms illustrate that healthy human adults prefer risky uncertainty condition to ambiguous uncertainty condition.

2 Previous theoretical approaches

Ambiguity aversion is called a paradox because it cannot be explained by the subjective expected utility theory or its variations like the prospect theory alone, which assumes that subjects make choices, with some errors, that maximize the expected utility, based on the judged probabilities of possible outcomes (Glimcher & Rustichini, 2004). In both the risky and ambiguous decks, the judged probability for one card to be either blue or red is exactly the same ($P_{\text{risk(Card=Red)}} = P_{\text{ambiguity(Card=Red)}} = 0.50$). One school of thought explains the ambiguity aversion in the following way: people use a heuristic that avoids betting when other people possess information that you lack, or when you lack information that would be helpful in making a decision. This explanation makes sense but is more like a situational moderator rather than a general explanation (Loewenstein et al., 2008).

Another school of thought allows choice to be influenced both by the best estimator (which is the mean) of subjectively judged probability (i.e., the degree of uncertainty) and the range of the subjectively judged probabilities (i.e., the degree of ambiguity). The spread or range of the
subjectively judged probabilities is larger in the ambiguous condition (varying from 0 to 1 with mean centered at 0.50) than in the risky condition (fixed at 0.50). If subjects are aversive to a wide spread of the judged probability even when the mean of the judged probability is the same, they show aversion to ambiguity (Becker & Brownson, 1964). Economists further applied the Minimax theorem to explain ambiguity: people are inclined to consider the worst possible outcome of each choice as the outcome that will occur and choose a set of choices that maximize the sum of the worst value of choices (Epstein, 2001; Rustichini, 2005). This is a widely accepted macro level economics model, and is developed to explain why international investors choose to put money in the stock market of a particular country (e.g., the U.S. or China). Countries are the sets, and stocks in each country are the choices.

The economy based theory has limited application to the experimental studies in ambiguity aversion in the two popular paradigms mentioned before. The economy theory only explains the preference between sets (ambiguous deck vs. risky deck). In each set, however, there are three choices (to bet on red, to bet on blue or to drop). The economy theory does not explain the preference among choices within a set. On the other hand, in order to apply Minimax theorem, economy theory has to figure out the worst value of choices over the long run in the stock market scenario. Unfortunately, it is difficult to determine how the worst value of choices in the risky deck is different from that in the ambiguous deck in the experimental paradigm. Should the worst values of a single choice of either red or blue card both be zero in both decks in the experimental paradigm?
Introduction

Imagine that a woman looking for anti-wrinkle cream can get only one free sample of either Crème de la Mer or Whoo’s Hwanyugo Cream (both of which retail for $200/bottle). She knows that (1) 100 samples of each cream have already been distributed and (2) 50 women reviewed Crème de la Mer favorably and 50 reviewed it negatively. No review is available yet for Hwanyugo cream. The best estimate of whether Hwanyugo cream works on wrinkles is 50:50, the same as Crème de la Mer. Thus, a woman should be indifferent (Hsu et al., 2005) in choosing the two types of cream samples according to the original expected utility theory, subjective expected utility theory (for their challenges in explaining ambiguity aversion, see Camerer & Weber, 1992), or prospect theory (Kahneman & Tversky, 1979; Tversky & Kahneman, 1992). Most women, however, will be more likely to choose Crème de la Mer than Hwanyugo Cream or more likely to re-sell Hwanyugo Cream than Crème de la Mer for $3 (drop option). This is an example of the empirical phenomenon called ambiguity aversion (Ellsberg, 1961). Frisch and Baron claimed “ambiguity is uncertainty about probability, created by missing information that is relevant and could be known” (cited in Camerer & Weber, 1992). Mathematically, ambiguity is “any distribution of probabilities other than a point estimate” (Becker & Brownson, 1964). Ambiguity and risk both describe an event that is probabilistic, and they together are called uncertainty. Previous work indicates most people are more averse to ambiguity than to risk (Camerer & Weber, 1992; Hsu et al., 2005). Models to explain this phenomenon depend upon mathematical sophistication such as integration over the second order probability distribution (Kahn & Sarin, 1988) or integration and axioms over probability distributions but without specified formula for the functions used in the model (Ghirardato, Maccheroni, & Marinacci,
Ambiguity aversion is closely related to a component of meta-cognition: uncertainty monitoring, which is defined as an animal knowing whether it feels uncertain (Crystal & Foote, 2009). To evaluate uncertainty monitoring, in repetitive trials, an objective stimulus is given to the animal, and the animal responds by classifying the stimulus as belonging to one of two categories. The animal receives a reward for a correct choice, no reward for an incorrect choice, and a lesser reward, compared to that of the correct choice, for “dropping” the trial if the drop option is present (decline paradigm). This paradigm is extremely similar to the drop option in Hsu’s (2005) experimental paradigm to evaluate ambiguity aversion in people. If we ask a subject to classify a stimulus previously associated with a given probability to high or low categories of probability and include a drop option, that provides an intermediate value of reinforcement, then this paradigm replicates the uncertainty monitoring paradigm. When the stimulus is associated with a probability of 0.5 compared to a probability near 1 or 0 (for example 0.8), we expect that subjects will drop the trial more often if they feel uncertain, and this would be evidence in support of meta-cognition so long as an associative explanation can be ruled out. However, there is a strong debate surrounding the evidence for meta-cognition in nonhuman animals. The stimulus response model presented by (Smith, Beran, Couchman, & Coutinho, 2008) predicts that the animal will drop more often when p=0.5 than when p is near the extreme values. But if the best estimate of probability is constant at p=0.5, the stimulus response model predicts risk aversion instead of ambiguity aversion. For example, the stimulus response model predicts subjects will drop trials more frequently when p=0.5 (risk) than when p = either 0.2 or 0.8 (ambiguity); see Appendix 1. In a similar situation, humans showed
ambiguity aversion (Halevy, 2007). Ambiguity aversion, as expressed by dropping trials more often when p is ambiguous than when p is risky if a drop option is presented, would indicate that animals feel uncertain when p is ambiguous. This outcome cannot be predicted by an existing associative model such as Smith et al’s stimulus response model, and thus would provide support for the presence of mega-cognition in animals.

A single published study addressed ambiguity aversion in a nonhuman species. Chimpanzees and bonobos more often chose a container previously shown to them to contain either good or bad food, in favor of a container with unknown contents (that is, the subjects did not see what was in the container until they had chosen it) (Rosati & Hare, 2011). We examined whether tufted capuchins (*Sapajus apella*) showed ambiguity aversion. Our design presented the choices in risky and ambiguous conditions in the same visual display. The subjects in this experiment avoided ambiguity, strengthening the evidence for continuity across primates in decision making under ambiguity, by eliminating the alternative explanation that the subjects in Rosati and Hare’s experiment never knew there was food in the ambiguous container.

Our Pan-Fragaszy Softmax model was originally developed to explain human ambiguity aversion inspired by Boltzeman softmax rule in decision making (Lee, 2006). This model produces major predictions like: Prediction 2: if the decision maker is averse to ambiguity and the probability levels expand the same range, then he/she will be more averse to discrete probability levels (special case is two probability levels) than a continuous uniform probability ranging from 0 (not including 0) to 1. This model has not been confirmed for humans or in nonhumans. Prediction 3: if the decision maker is averse to ambiguity and sensitive to the magnitude of ambiguity, then he/she will be more averse to two probability levels spread out at larger range than at smaller range (i.e., more aversion to p=0.2 or 0.8 than p=0.4 or 0.6);
The aim of our study was to see whether capuchin monkeys chose differently in ambiguous condition vs risky conditions, and particularly, whether they were averse to ambiguity, as humans are.

**Method**

Three monkeys learned in a training phase to choose either a green square or a black square (two alternative choices) to match a stimulus square. Matching the color always delivered 5 pellets while a non-matching choice never delivered a pellet. Then they learned that a drop circle delivers one pellet (for two monkeys) and two pellets for the other monkey (Appendix 3 result). Then they were tested on 3 ambiguous conditions, 1 risky condition and 1 certain condition: (1) Ambiguity-Low: p=0.4 or 0.6, (2) Ambiguity-High: p=0.2 or 0.8, (3) Ambiguity-Range: p=0, 0.2, 0.4, 0.6, 0.8 or 1, (4) Risk: p=0.5, and (5) Certain: same color matching rule as in the training phase. The Risky conditions were interspersed between Ambiguous conditions, and the order of presentation of ambiguous conditions was balanced across subjects. In all the conditions, they had two choices (a green square and a black square). The green square delivered 5 pellets with the assigned probability and a drop option (a purple circle) constantly delivered less than 5 pellets (Figure 1 Testing conditions). When p=0.4, 40% of the time, choosing the green square delivered 5 pellets, and 60% of the time choosing the black square delivered 5 pellets. The IACUC approval of animal use protocol number is AUP #: A2012 02-026-Y1-A0. See Appendix 2 for details of method. We predicted that the monkeys would choose the drop option differently in the risk condition compared to any of the three ambiguous conditions, as well as from the Certain condition. If they have aversion to ambiguity, they would drop less in the risky condition than in any of the ambiguous conditions.
Figure 1. The square in the middle of each screen is stimulus square. The black, green squares are the two choices with a potential to deliver 5 pellets and purple circle is drop.
Results

After controlling for individual differences, three monkeys showed aversion to Ambiguity-High and Ambiguity-Low conditions (i.e., they dropped more often in the Ambiguity-High and Ambiguity-Low conditions compared to Risk conditions), dropped significantly less in Certain condition, and dropped almost significantly less in Ambiguity-Range condition all compared to Risk condition (table 1). There is a significant individual difference between Xenon and the other two monkeys (Appendix Table 1 in Appendix 3). Xenon dropped significantly more often in Ambiguity-Range condition compared to Risk condition (Appendix Table 9 and table 2). The other two monkeys dropped more often but not significantly so in Ambiguity-Range condition compared to Risk condition. (Appendix Table 7, 8 and table 2). There is no significant difference in the number correct choices vs. the number of incorrect choices among any uncertain conditions ($\chi^2$ (d.f.=3) = 1.871, $p = 0.600$) (Appendix Table 10).

Table 1. Logistic regression of 3600 trials from three monkeys.

<table>
<thead>
<tr>
<th>Conditions</th>
<th>B</th>
<th>S.E.</th>
<th>Wald</th>
<th>df</th>
<th>Sig.</th>
<th>Exp(B)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ambiguity-Range vs.</td>
<td>-.334</td>
<td>.183</td>
<td>3.313</td>
<td>1</td>
<td>.069</td>
<td>.716</td>
</tr>
<tr>
<td>Risk</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Certain vs. Risk</td>
<td>-.883</td>
<td>.208</td>
<td>17.972</td>
<td>1</td>
<td>&lt;.0001</td>
<td>.414</td>
</tr>
<tr>
<td>Ambiguity-High vs.</td>
<td>1.202</td>
<td>.154</td>
<td>60.992</td>
<td>1</td>
<td>&lt;.0001</td>
<td>3.325</td>
</tr>
<tr>
<td>Risk</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ambiguity-Low vs.</td>
<td>.743</td>
<td>.158</td>
<td>22.215</td>
<td>1</td>
<td>&lt;.0001</td>
<td>2.103</td>
</tr>
<tr>
<td>Risk</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2. drop option 1 or 2 pellets (vs 5 for correct choice)

<table>
<thead>
<tr>
<th>Number of dropped trials in each condition</th>
<th>240 trials per monkey per condition</th>
<th>sum</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Leo</td>
<td>Chris</td>
</tr>
<tr>
<td>Risk</td>
<td>10</td>
<td>13</td>
</tr>
<tr>
<td>Ambiguity-Range</td>
<td>15</td>
<td>22</td>
</tr>
<tr>
<td>Ambiguity-High</td>
<td>28</td>
<td>27</td>
</tr>
<tr>
<td>Ambiguity-Low</td>
<td>40</td>
<td>20</td>
</tr>
<tr>
<td>Certain</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>Risk (Phase 5)</td>
<td>14</td>
<td>N.A.</td>
</tr>
</tbody>
</table>

**Discussion**

Our results indicate that capuchins present ambiguity aversion as described for humans (Halevy, 2007) and chimpanzees and bonobos (Rosati & Hare, 2011) as well as supporting uncertainty monitoring. We found capuchins were averse to Ambiguity-Low (p=0.4 or 0.6) and Ambiguity-High (p=0.2 or 0.8) compared to Risk (p=0.5). The effect size, as indicated by exponentiation of the “B” value (Burns & Burns, 2008 p.583), of Ambiguity-High is larger than Ambiguity-Low as was predicted by the Pan-Fragaszy Softmax model. It’s unclear about capuchin’s preference to Ambiguity-Range compared to Risk.

**Pan-Fragaszy Softmax model**

The softmax rule describes that the choice is determined probabilistically on the basis of the actions’ expected values. The probability of choosing a particular option is given by the
Boltzmann distribution of the value functions (Lee, 2006). When the option’s value is probabilistically determined, the decision maker will choose based on the expected utilities of the options. The subjective expected utility of the \(j^{th}\) choice is

\[
\hat{\mu}_j = p_j v_j, \tag{1}
\]

in which \(p_j\) is the subjectively judged probability of the option \(j\) yielding a value \(v_j\). To simplify the presentation of our model, we use objective probability and objective value in place of subjective probability and subjective value. The model permits subjective weighting of objective probability and subjective value function of objective value to be used in equation (1) to calculate subjective expected utility, as in prospect theory (Kahneman & Tversky, 1979).

We propose, if the subjective expected utility of the \(j^{th}\) option is \(\hat{\mu}_j > 0\), the probability of choosing the \(j^{th}\) option (denoted by \(P_j\)) is

\[
P_j \propto e^{-\beta/\hat{\mu}_j}. \tag{2}
\]

where \(\beta\) is a positive real parameter to indicate individual attitude, and is called attitude parameter. The function that \(P_j\) is proportional to, or \(e^{-\beta/\hat{\mu}_j}\), is called decision dependence function of the \(j^{th}\) option, or \(D(p_j)\), because it is the exponential dependence of the probability of choosing an option based on the subjective expected utility of that option.

\[
P_j = \frac{e^{-\beta/\hat{\mu}_j}}{\sum_{j=1}^{n} e^{-\beta/\hat{\mu}_j}}, \quad \text{given that } \hat{\mu}_j > 0, \text{ for all } j \in [1,n] \tag{2}
\]

in which the sum of \(e^{-\beta/\hat{\mu}_j}\) over all \(n\) options is the normalizing constant to guarantee that

\[
\sum_{j=1}^{n} P_j = 1.
\]

From equation (1), (2) and (3), the probability \(P_j\) of choosing the \(j^{th}\) option from \(n\) options of positive yield (i.e., \(v_j > 0\),)
\[ P_j = \frac{e^{-\beta/(p_j v_j)}}{\sum_{j=1}^{n} e^{-\beta/(p_j v_j)}}, \quad \text{given that } v_j > 0, \text{ for all } j \in [1,n] \quad (4). \]

We propose that when the jth choice is ambiguous, the probability of choosing the jth choice is proportional to the expected value of the decision dependence. When the ambiguous probability distribution of the jth option forms a discrete uniform distribution taking on k probability values: \( p_{j1}, p_{j2}, p_{j3}, \ldots, p_{jk} \), the expected value of the decision dependence of the jth option is:

\[ P_j \propto E(D(p_j)) = \frac{1}{k} \left( e^{-\beta/(p_{j1} v_j)} + e^{-\beta/(p_{j2} v_j)} + \cdots + e^{-\beta/(p_{jk} v_j)} \right) \quad (5). \]

When the ambiguous probability of the jth option is continuous, the expected value of decision dependency as a function of \( p_j \), \( D(p_j) = e^{-\beta/(p_j v_j)} \), with respect to the probability density function (p.d.f.) \( \Phi(p_j) \), is given by the inner product of \( \Phi(p_j) \) and decision dependency:

\[ P_j \propto E \left( D(p_j) \right) = \int D(p_j) \Phi(p_j) \, dp_j \quad (6). \]

Note that in most experimental paradigms, the risky option yields the same value as the ambiguous option, i.e., \( v_j = v \).

If the ambiguous probability of the jth option is continuously uniform on \((0,1)\), then the p.d.f. of \( \Phi(p_j) = 1 \). The expected value of the decision dependency of the jth option is:

\[ P_j \propto E \left( D(p_j) \right) = \int_0^1 e^{-\beta/(p_j v_j)} \, dp_j = \int_0^1 e^{-\beta/p} \, dp_j = e^{-\beta/v} - \frac{\beta \Gamma(0,\beta/v)}{v}, \text{ when } \frac{\beta}{v} \geq 0. \quad (7) \]

If the ambiguous probability of the jth option is uniformly continuous on \([a_j,b_j]\), we are inclined to change the probability density of \( p_j \) to \( f(p_j) = 1/(b-a) \) to keep the cumulative probability of \( \Phi(p_j) \) equal to 1.
P_i \propto E(D(p_j)) = \int_{b_j}^{a_j} e^{-\frac{\beta}{p_j}} * \frac{1}{b_j-a_j} * d(p_j) = -a_j e^{-\frac{\beta}{\bar{p}c}} + b_j e^{-\frac{\beta}{\bar{b}c}} + \frac{\beta \Gamma(0, \frac{\beta}{\bar{p}c})}{\nu}

when 0 \leq a_j < b_j < 1, and a_j and b_j are real numbers. \hspace{1cm} (8)

The Pan-Fragaszy Soft model predicts ambiguity aversion when \( \beta \) is small \( \text{ (for a special case, see Appendix 4)} \). We are curious about the range of \( \beta \) that allows ambiguity aversion to happen in the classical Ellsberg’s urn problem, where we interpret the ambiguity as the choosing the ambiguous option (ambiguous urn) will yield value V with an ambiguous probability that an uniform continuous distribution over (0,1). An analytical solution to find the range of \( \beta \) can be achieved by solving the inequality:

\[ P_{l(\text{risk})} = D(p) = e^{-\frac{\beta}{\bar{p}c}} > \int_0^1 e^{-\frac{\beta}{p}} d(p) = e^{-\frac{\bar{p}c}{\nu}} - \frac{\beta \Gamma(0, \frac{\bar{p}c}{\nu})}{\nu} = D(p) = P_{m(\text{ambiguous})} \]

with respect to \( \beta \). \( \text{(9)} \)

The above inequality does not have an analytical solution with respect to \( \beta \). It can be solved numerically by plotting both side of the inequality on a graph and find the range of \( \beta \) with given \( \nu \) values. Thus we do not pursue the analytical solution but instead limit our further discussion to the conceptual shape of the integrand. Let \( a = \frac{\beta}{\nu}, x=p_j \), then \( y = e^{-ax} \), which is the generic form of the decision dependency function. Figure 3 shows the curve of \( y = e^{-ax} \) with different values of \( a \).
The properties of the integrand $y = e^{-a/x}$:

1. $y = e^{-a/x}$ is monotonically increasing, the derivative of the integrand is:
   
   $a e^{-a/x} > 0$, when $x > 0$, $a > 0$, and $x$ and $a$ are real numbers.

2. It changes from a concave upward to convex upward at $x = \frac{b}{2a}$. When $x$ varies from 0 to $\infty$. The second derivative of $y = e^{-a/x}$ is:

   
   $\frac{b^2 e^{-b/a} - 2be^{-b/ax}}{a^2 x^4} = \frac{2be^{-b/ax}}{ax^3}$

   Set the second derivative equal to 0, we get
   
   $x = \frac{b}{2a}$.

---

Footnote: Concave upward is more frequently called convex in economic literature and convex upward is called concave.
(3) The limit of $e^{-a/x} = 1$ when $x \to \infty$.

Predictions of Pan-Fragaszy Softmax model:

(1) Most normal people are more averse to ambiguity than to risk; people with certain brain injuries may be more averse to risk than to ambiguity (Hsu et al., 2005). In our model, this attitude depends on the value of $\beta$. People with $\beta$ that makes the curve of $y$ mostly convex upward at $x \in (0,1]$ will be ambiguity averse and people with $\beta$ that makes the curve of $y$ mostly concave upward at $x \in (0,1]$ will be risk averse (Figure 3). The integral of $e^{-a/p}$ with respect to $p$ is:

$$e^{-\frac{a}{p}}p + a \text{ExpIntegralEi}\left(-\frac{a}{p}\right).$$

Cumulative distribution function of decision dependency is:

$$\int_0^pe^{-\frac{\beta}{vp}}d(p) = pe^{-\frac{\beta}{vp}} - \frac{\beta \text{Gamma}[0, \frac{\beta}{vp}]}{v}$$ (10)

Mass function of decision dependency on range $(c,d)$ $(0<c<d \leq 1)$ is:

$$\int_c^d e^{-\frac{\beta}{vp}}d(p) = -ce^{-\frac{\beta}{vc}} + de^{-\frac{\beta}{vd}} + \frac{\beta \text{Gamma}[0, \frac{\beta}{vc}]}{v} - \frac{\beta \text{Gamma}[0, \frac{\beta}{vd}]}{v}$$ (11)

Expected value of decision dependency on $(c,d)$ with p.d.f. $\Phi(pj)=1/(d-c)$ is:

$$\frac{1}{(d-c)} \ast \{-ce^{-\frac{\beta}{vc}} + de^{-\frac{\beta}{vd}} + \frac{\beta \text{Gamma}[0, \frac{\beta}{vc}]}{v} - \frac{\beta \text{Gamma}[0, \frac{\beta}{vd}]}{v}\}$$ (12)

A general condition that satisfies ambiguity aversion with respect to $\beta$, where risk $p= (c+d)/2$ and ambiguous probability p.d.f. is $\Phi(pj)=1/(d-c)$ on $(c,d)$ is:

$$e^{-\frac{\beta}{vc+c+d}} > \frac{1}{(d-c)} \ast \{-ce^{-\frac{\beta}{vc}} + de^{-\frac{\beta}{vd}} + \frac{\beta \text{Gamma}[0, \frac{\beta}{vc}]}{v} - \frac{\beta \text{Gamma}[0, \frac{\beta}{vd}]}{v}\}$$ (13)

\[2\text{ For more information on ExpIntegralEi, please refer to the webpage: http://mathworld.wolfram.com/ExponentialIntegral.html]
In the original problem presented by Ellsberg (1961), the ambiguous urn’s c is very near to 0, d=1, so that inequality (13) should be written as inequality (9).

Figure 3 shows ambiguity aversion in a special case with β=0.2, c=0.2, and d=0.8. Because the expected value of ambiguous decision dependency is smaller than the risky decision dependency, the subject is predicted to show ambiguity aversion.

(2) If the decision maker is averse to ambiguity, then he/she will be more averse to discrete probability levels (special case is two probability levels) than a continuous uniform probability ranging from 0 (not including 0) to 1. This prediction has not been tested for humans or in nonhumans.
(3) If the decision maker is averse to ambiguity and sensitive to the magnitude of ambiguity, then he/she will be more averse to two probability levels spread out at larger range than at smaller range (i.e., more aversion to p=0.2 or 0.8 than p=0.4 or 0.6 or more averse to p from 0.2 to 0.8 than p from 0.4 to 0.6). This is confirmed in our results. The effect size of Ambiguity-High condition is larger than the effect size of Ambiguity-Low. This prediction was explored by Becker and Brownson (1964) in humans but these authors did not apply inferential statistics to their data.

(4) The monetary value of an option (i.e., if value for correct choice =$10 and drop = $2 vs. correct = $100 and drop = $20) does not change the preference. This is supported in humans (Halevy, 2007).

(5) If the subjective probability is centered at \( p_0 \in (0,1) \) with a symmetric range of length L at either side of the center, our model predicts that ambiguity aversion will occur if the range of the judged probability (x) is more toward the convex upward part of the \( y= e^{-ax} \) curve than toward the concave upward part of the curve (see figure 3. That is, when L is held constant, the person is more likely to show ambiguity aversion when \( p_0 \) is near 1 than when \( p_0 \) is near 0.

The prevailing models concerning choice under conditions of risk and ambiguity come from a different approach. The most influential model that explains ambiguity aversion in humans is maxmin expected utility theory (MEU) (Gilboa & Schmeidler, 1989), recursive expected utility (KMM) (Klibanoff et al., 2005) and GMM (Ghirardato et al., 2004). MEU is a special case of KMM and GMM. KMM and GMM have terms for pure risk (single point probability) and ambiguity (second order probability distributions). Like our model, predictions from KMM or GMM are not in conflict with those from EU, SEU or prospect theory. Any utility functions EU, SEU or Prospect theorey use, can be put into the risk term of KMM or GMM.
Unlike our model, the shape of the function defining the ambiguity attitude is not specified by a formula in KMM or GMM. Other models (including KMM) are not perfect in humans (Halevy, 2007). Kahn and Sarin attempted to specify a formula for ambiguity aversion (Kahn & Sarin, 1988). However, their model might be problematic because if we rewrite their formula to have separate risk term and ambiguity term, we can see that the risk component is weighted to have a value less than zero when \( p \) is small and the weight of pure risk is convex upward, which is in conflict with the confirmed predictions of prospect theory’s weight toward pure risk probability (Appendix 5).
Conclusions

Pan-Softmax model is a new behavioral economic model with specified formula and predictions on choice concerning risk and ambiguity. It partially explained human and non-human’s responses to ambiguity. It’s not perfect yet, but its attempt to specify formula will lead to clearer testable predictions for future studies than models rely on axioms.

The findings that capuchins are averse to ambiguity provide much stronger evidence for meta-cognition in any non-human animals than previously available. The previous evidence for meta-cognition relies on uncertainty monitoring by classifying the stimulus as belonging to one of two categories in a drop paradigm. This study addressed uncertainty monitoring from an innovative approach that is different from any previous variations of uncertainty monitoring task. In the previous variations of uncertainty monitoring tasks, the objective stimulus took one value at a time. In our study, the objective probability in the ambiguous condition took multiple values and a time. The results we got cannot be predicted by any models from associative learning’s point of view.
Appendices

Appendix 1 Stimulus response model cannot predict ambiguity aversion

The stimulus response model proposed by (Smith et al., 2008) proposed a model for why animals drop for a lesser reward when they can choose one of the two better rewarded alternative choices. The model predicts more frequent uncertain responses (i.e., drop for a lesser amount of reward) when the stimulus is in the middle of two the criteria for the two alternative choices than when the stimulus is away from the middle (Smith et al., 2008).

In Appendix Fig.1, the objective probability is the objective stimulus in Smith et al. (2008)’s stimulus response model; the subjective probability levels is analogous to subjective levels in the stimulus response model; the secondary probability density (formed in the subject’s perception) of an objective probabilistic stimulus (Y axis in Appendix Fig. 1) is analogous to the probability density in the stimulus response model. The criterion lines for a subject to drop are line AC and line BD in Figure 1. The perception of a risky stimuli at p=0.5 (risky) is depicted as the large solid bell shaped curve in Figure 1. The response strength for the subject to drop is represented by the area under the bell shaped curves between [C, D], or area ACDBI. First, if an ambiguous objective stimulus with a mean $\bar{p} = 0.5$ is interpreted as the probability of either 0.25 or 0.75 (set of probabilities), the perception is depicted by two skewed and truncated dash line curves with expected value of 0.25 and 0.75. The response strength for a subject to choose the decline option is the sum of area ECD plus area DFC. Because the area under the big solid curve (which is 1) equals the sum of the area under the two small dash line curves, area ACDBI is larger than the sum of area ECD and area DFC. According to the stimulus response model, this indicates that the subject is more likely to drop when the objective probabilistic stimulus is risky (p=0.5) than when it is ambiguous (p=0.25 or p=0.75). Second, if an ambiguous objective stimulus is
interpreted as a probability uniformly spreading between 0 and 1 (range of probabilities), it is
depicted as a horizontal dash line at probability density y =1 in Figure 1. The response strength
to the decline option is area ACDB. Still, area ACDB is smaller than area ACDBI. This indicates
that the subject is more likely to decline a trial when a risky stimulus is given than when an
ambiguous stimulus is given.

Appendix Figure 1. Stimulus response model predicts that animals are more likely to drop a trial
for a lesser amount of reward when the option in the trial is risky than when it is ambiguous.
Appendix 2 Appendix Method

Subjects and housing

Subjects were six adult male tufted capuchin monkeys (Sapajus apella), 19–23 years old in 2011, at the University of Georgia. Subjects are pair-housed with the rest of the four male capuchins in the colony. They volunteered to enter a transport box to be carried from their housing room to an adjacent testing room no more than once per day. None of them were food-deprived or water-deprived during the course of testing. Each test session lasted no more than one hour. Each monkey was tested individually in front of a computer monitor (Appendix Figure 2). Usually there were two monkeys from the same home cage tested concurrently on two separate computers in the testing room. The two monkeys being tested concurrently were in testing cages oriented in opposite directions, so when working at the task, the monkeys faced away from each other. A water bottle was attached to the side of each testing cage. If a monkey detached the water bottle from the designated place on the side testing cage, the water bottle would be placed on top of the testing cage and it would still have free access to water. All six monkeys had extensive prior experience in moving a virtual object (designated as a cursor) controlled by a joystick to another object (designated as a goal) on a computer screen. Only three monkeys (Leo, Chris and Xenon) met criteria for the final testing phase.

Test apparatus

Testing was automated by the Matlab program. Visual stimuli were presented on a Dell standard 12 inch (30.5 cm diagonal) LCD desk top computer monitor with 1024*768 resolution. A Logitech 3D force joystick was mounted 20 cm below the monitor and a Med Associate’s 20 MG Pellet via a National Instrument 6501 USB board were connected to the computer. Audio
signals were given by the speakers of the computer. A monkey stayed in a testing cage in front of the monitor during the experiment and it could reach to the joystick and pellet receptacle through a round aperture in the front plexiglass panel of the testing cage (Appendix Figure 2).

Appendix Figure 2. Test apparatus. Monkey Leo’s left hand was moving the joystick. The half-cylinder shaped plexiglass receptacle in the lower left side of figure was where pellets were delivered.

*Task presentation*

A non-visible X (from left to right) axis divided the screen into 400 units. A non-visible Y axis (from bottom to top) divided the screen into 300 units. The aspect ratio of the X axis to Y axis was 1:1, i.e., a unit in the Y axis equaled a unit in the X axis. First, a grey curser with 10 unit arm length and 10 point arm width appeared at the point (200, 250) and a sample (stimulus)
square with 60 unit side appeared in the center of display, at the point (200, 150) (Appendix Figure 3). The cursor could be moved by the joystick at up to 50 units per second along one axis, and the maximum speed was achieved when the joystick is moved completely to one side of the joystick’s base. Upon the cursor reaching the sample square, one or two choice square(s) of exactly the same shape as the sample square (in all phases) and a purple “drop” circle that indicates the drop option with 30 unit radius (excluding Phase 1) appeared on other parts of the screen as specified below (Appendix Figure 4). The cursor could continue to be moved, and the sample square could be moved in synchrony with the cursor through Phase 1.2.2. Third, if the cursor reached a choice square and if the choice square was designated as correct, an audio signal was played and five pellets were dispensed to the monkey. If the choice square was designated as incorrect, the second audio signal was played but no pellet was dispensed. If the cursor reached the dropping circle, the third audio signal was played and one pellet was dispensed. Then the program proceeded to the next trial. There was no inter-trial interval. If the cursor did not reach a choice square or a drop circle in 3 minutes, the program proceeded to the next trial.
### Testing phases

Appendix Table 1. The overall design of all phases.

<table>
<thead>
<tr>
<th>Phase</th>
<th>Aim</th>
<th>Stimulus (sample)</th>
<th>Options</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Learn match to sample</td>
<td>Certain</td>
<td>Matching= 5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Non-matching = 0</td>
</tr>
<tr>
<td>2</td>
<td>Learn drop circle</td>
<td></td>
<td>Matching = 5 or Non-matching = 0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Drop = 1</td>
</tr>
<tr>
<td>3</td>
<td>Control drop circle</td>
<td>Uncertain</td>
<td>Drop = 1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Null = 0</td>
</tr>
<tr>
<td>4</td>
<td>Testing</td>
<td>Certain and uncertain</td>
<td>Drop=1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Green=p*5</td>
</tr>
<tr>
<td>5</td>
<td>Control trial number</td>
<td></td>
<td>Black=(1-p)*5</td>
</tr>
</tbody>
</table>

1: Certain and uncertain stimulus are displayed in figure 1 and described in phase 4.

2: p in phase 4 is described in phase 4.

---

Phase 1: Match to sample color training.

In phase 1, the subjects learned to match the sample square from one matching and one non-matching choice square located equidistantly from the sample square. In phase I, each session consists of 40 initial trials. In each initial trial, the sample square had a 50:50 chance of pseudo-randomly being either green or black. There were two choice squares, one black and one green. They were located on the same horizontal line as the samples square, i.e., the Y coordinates of the centers of all three squares were 150. If the sample was black, then the black choice square
was the matching choice, and vice versa. The black choice square had a 50:50 chance to be pseudo-randomly located on the left or right side of the sample while the opposite side is the green choice square and vice versa. At the beginning of each session, a new and unique seed was calculated for the generation of the pseudo-random numbers. The pseudo-random numbers for one purpose (e.g., colors) were different from another (e.g., the location) but the seed for pseudo-random number generator was the same for all purposes. The Mersenne Twister with improved initialization method was used to generate pseudo-random numbers, and this method was the same throughout all phases of the experiment.

Appendix Figure 3 phase1
Phase 1.1: Match to sample color training, unequal distance.

In phase 1.1, when the choice squares appeared, the distance from the matching choice square to the sample was shorter than the distance from the non-matching choice to the sample. There are three levels (phase 1.1.1, 1.1.2, 1.1.3) of distance combinations (Appendix Table 2). If the monkey's initial responses were correct at least 80% of the initial trials for two consecutive sessions, the monkey moved on to the next level (hereafter, criterion 1).

Appendix Table 2. Phase 1 distances from options to the sample in units

<table>
<thead>
<tr>
<th>Phase</th>
<th>matching square</th>
<th>non-matching square</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1.1</td>
<td>75</td>
<td>150</td>
</tr>
<tr>
<td>1.1.2</td>
<td>100</td>
<td>150</td>
</tr>
<tr>
<td>1.1.3</td>
<td>130 (Leo 135)</td>
<td>150</td>
</tr>
<tr>
<td>1.2</td>
<td>135</td>
<td>135</td>
</tr>
</tbody>
</table>

In phase 1.1, once the cursor reaches the sample square, the sample square would move in synchrony with the cursor. Moving the cursor to the matching choice square was the correct response. If the response was correct, 5 pellets are dispensed, an audio signal was given and the program proceeded to the next initial trial. If the response was incorrect, a correction procedure was applied. A correction trial repeated the display of the immediate past initial trial. Correction trials repeated until the monkey chooses correctly or until three minutes pass. Then the program displayed the next initial trial. Only the choices in the initial trials (i.e., the first presentation) were calculated for the data analysis and criterion meeting. This correction procedure effectively prevented side bias.
Phase 1.2: Match to sample training, equal distance.

In phase 1.2, the distance from the sample to the matching and non-matching squares was the same (Appendix Table 1). The rest of the procedures in phase 1.2.1 were the same as those in Phase 1.1.1. A monkey proceeded to Phase 1.2.2 after meeting criterion 1. In phase 1.2.2, the sample no longer moved with the cursor after the cursor reaches the sample. The rest of the procedures and criterion in Phase 1.2.2 were the same as those in Phase 1.2.1. In phase 1.2.3, the choice squares could be located in two of three locations of the same set. The first set of three potential locations was: to the right of the sample, to the top left of the sample, and to the bottom left of the sample. The second set of the three potential locations was the mirror locations of the first set: to the left of the sample, to the top right of the sample, and to the bottom right of the sample. Within a set, the angular distance between either two of the potential locations to the center of the sample was the same, $\frac{2\pi}{3}$; the Euclidean distance from any one of the potential location to the center of the sample was 135 units$^3$. In a trial, the use of one set or another was determined pseudo-randomly at a 50:50 chance. A monkey needed to make at least 80% percent of the responses correctly per session, at least 36 trials per session for 4 consecutive sessions (hereafter, criterion 2) to move to Phase 2.

Phase 2: Training to learn the value of a drop option

In phase 2, monkeys learned that the decline option, a purple dropping circle, gives them one pellet in a trial. There were 40 initial trials per session. The sample square was generated the same way as in Phase 1. The location of the choice square and the dropping circle was the same as those in Phase 1.2.3. Correction procedure as described in phase 1.1 was applied.

---

$^3$For some monkeys in some sessions, the distance is 130 units. The units are adjusted for practical training of individual monkeys.
Phase 2

Appendix Figure 4 phase2

Phase 2(0): Dropping vs. non-matching choice.

Only Leo was trained with Phase 2(0). The centers of dropping circle and the non-matching choice square randomly appeared in two of the three potential locations from one set around the sample (Appendix Figure 5). Because choosing the dropping circle yields 1 pellet but nothing for non-matching choice square, moving the cursor to the dropping circle was the correct response. Leo proceeded to Phase 2 after meeting criterion 1.

Phase 2 (main): Dropping vs. matching or non-matching choices.

Because Leo developed a bias to choose the drop option, the other two monkeys began their phase 2 training directly from phase 2 (main). The centers of dropping circle and the choice square randomly appeared in the three locations in one set. Of the 40 initial trials in a session, a randomly ordered 20 trials had a matching choice square and a dropping circle (Figure 6); the other 20 trials had a non-matching choice square and a dropping circle. When a matching choice square was presented, moving the cursor to the match choice square was the correct choice and produced 5 pellets while moving the cursor to the dropping circle was the incorrect choice and
produced 1 pellet. In the case of a non-matching choice square presented, the designation of the correct response was the same as those described in Phase 2(0). A monkey proceeded to Phase 3 after meeting criterion 2.

Phase 3: Control drop option.

The display of a trial was similar to that in phase 2, except (1) that the stimulus was one of the four uncertain stimuli: Risky, Ambiguity-Range, Ambiguity-High and Ambiguity-Low as described in phase 4 and (2) that the two choices were the dropping circle with a 1 pellet reward and a yellow triangle with no reward (Appendix Figure 5). Within each 4-trial unit, the order of the four uncertain stimuli was random. In each session, there were 10 repetitions of the 4-trial unit. Each monkey did 4 sessions (160 total trials) or 40 trials of each stimulus. There was no correction procedure from Phase 3 and on. From Phase 3 and on if a monkey did not complete a initial trial within three minutes, the following trial would be the immediate past trial until the monkey completed this initial trial; that is, the monkey would not skip any initial trial.
Phase 4: Testing.

There were five types of sample squares in Phase 4: Certain and 4 uncertain types (Risky, Ambiguity-Range, Ambiguity-Low, and Ambiguity-High) (Figure 1). When the cursor reaches the sample square, the two choice squares (one black and one green) and the dropping circle, randomly appeared on each of the three locations, as described in Phase 1.2.3. In all the conditions, the monkeys had two choices (a green square and a black square). The green square delivered 5 pellets with the assigned probability and a drop option (a purple circle) constantly
delivered less than 5 pellets (Figure 1 Testing conditions). When p=0.4, 40% of the time, choosing the green square delivered 5 pellets, and 60% of the time choosing the black square delivered 5 pellets.

In a trial with reward probability of 1 if a matching choice was chosen, the sample square was randomly and equally likely to be either black or green. The matching choice square was the correct choice and the non-matching choice square was the incorrect choice. This is the Certain condition.

In a trial with p=0.5 to reward choosing the green square, the sample square was half green on the right side, and half black on the left side. The program pseudo-randomly selected either a green or black choice square to be the correct choice square at a 50:50 chance level, i.e., choosing either one of the choice square yields five pellets 50% of the time. This is the Risky condition.

In a trial in the Ambiguity-Range condition, the reward probability for choosing the green square was p=0, 0.2, 0.4, 0.6, 0.8 or 1. The stimulus square was a shade gradient from black on the left to green on the right. The reward schedule for each monkey in Ambiguity-Range condition is described in Appendix Table 3. A list of whether a green choice was designated as correct in each trial was predetermined precisely according to the reward probability, instead of generated randomly.
### Appendix Table 3. Reward probability level for ambiguous trials in Phase 4.

<table>
<thead>
<tr>
<th>Monkey</th>
<th>Conditions</th>
<th>40 trial sessions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Leo</td>
<td>Ambiguity-Range (0, 0.2, .4, .6, .8, 1)</td>
<td>0.8</td>
</tr>
<tr>
<td></td>
<td>Ambiguity-High (0.2 vs. 0.8)</td>
<td>0.8</td>
</tr>
<tr>
<td></td>
<td>Ambiguity-Low (0.4 vs. 0.6)</td>
<td>0.4</td>
</tr>
<tr>
<td>Xenon</td>
<td>Ambiguity-Range (0, 0.2, .4, .6, .8, 1)</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td>Ambiguity-High (0.2 vs. 0.8)</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td>Ambiguity-Low (0.4 vs. 0.6)</td>
<td>0.4</td>
</tr>
<tr>
<td>Chris</td>
<td>Ambiguity-Range (0, 0.2, .4, .6, .8, 1)</td>
<td>0.6</td>
</tr>
<tr>
<td></td>
<td>Ambiguity-High (0.2 vs. 0.8)</td>
<td>0.8</td>
</tr>
<tr>
<td></td>
<td>Ambiguity-Low (0.4 vs. 0.6)</td>
<td>0.6</td>
</tr>
</tbody>
</table>

In a trial in the Ambiguity-Low condition, two identically shaped sample squares alternate at the center of the screen twice per second. One sample square consists of 40% of black on the left and 60% of green on the right; the other sample square consists of 40% of black on the left and 60% of green on the right. The probability for reward choosing the green square was 0.4 or 0.6, as detailed in Appendix Table 3.

In a trial in the Ambiguity-High condition, two identically shaped sample squares alternate at the center of the screen twice per second. One sample square consists of 20% of black on the left and 80% of green on the right; the other sample square consists of 20% of black on the left and 80% of green on the right. The probability for reward choosing the green square was 0.2 or 0.8, as detailed in Appendix Table 3.
Trials of the Certain condition were presented at the beginning of Phase 4. Trials of the Risky conditions were interspersed between trials of the three ambiguous conditions. The order of ambiguous conditions was balanced across subjects. The monkeys were tested in the order described in Appendix Table 3. A monkey was tested until one hour had passed or until it did not respond for 10 minutes, whichever happened first.

**Appendix Table 4. Order of types of trials for each monkey.**

<table>
<thead>
<tr>
<th>Monkey</th>
<th>Conditions†</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>C</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Leo</td>
<td>C</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Chris</td>
<td>C</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Xenon</td>
<td>C</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Number of trials | 240 | 60 | 40*6=240 | 60 | 40*6=240 | 60 | 40*6=240 | 60 |

†: C=Certain, R=Risky.

Phase 5: Control trial number per session.

Leo and Xenon finished about 100 or more trials in the ambiguous conditions within the one hour testing time but only fixed 60 trials in the Risky conditions in Phase 4. Phase 5 was added to avoid an alternative explanation that any difference in dropping a trial between an ambiguous condition and the risky condition was due to fewer trials per session in the risky
condition. Phase 5 presented 120 trials per session of the Risky condition. Everything else in Phase 5 was the same as in Phase 4. Only Leo and Xenon did Phase 5 because Chris did not do significantly more trials in the ambiguous conditions in Phase 4 (mean trials per session in Risk condition (n=7): 18.74, mean trials per session in ambiguous conditions (n=14): 23.84, \( t = -1.656 \), \( p = 0.11 \), independent two sample t-test).

Appendix 3 Appendix results

Appendix Figure 6 shows the learning curve for each monkey in Phase 1 and Phase 2. Appendix Table 5 shows that there was no difference in the rate of dropping a trial among all the uncertain conditions, i.e., any difference in dropping a trial among all the uncertain conditions found in phase 4 was due to different reward probability levels associated with the stimulus, not the visual properties of the stimulus itself.
Appendix Figure 6. Learning curves in Phase 1 and Phase 2.
Appendix Table 5. Number of dropped trial in each condition of 40 trials in Phase 3.

<table>
<thead>
<tr>
<th>Monkey</th>
<th>Ambiguity-High</th>
<th>Ambiguity-Low</th>
<th>Risk</th>
<th>Ambiguity-Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leo</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Chris</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Xenon</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Separate logistic regressions were run with all data from Phase 4 and each monkey separately. The dependent variable was whether the monkey dropped a trial. For all monkeys, the independent variable of interest was the conditions. The independent variables not of interest were: individual difference and location of the drop circle (six possible locations as described in Phase 3). Individual difference was redundant with whether dropping a trial delivered 1 or 2 pellets. For each monkey, the individual difference was taken out of the regression model. Only Xenon dropped significantly more often in the Risky condition than in the Ambiguity-Range condition. There was no difference for any monkey in the rate of correct choices among uncertain conditions in Phase 4 and Phase 5 (Appendix Table 10).
### Appendix Table 6. Logistic regression for three monkeys

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>S.E.</th>
<th>Wald</th>
<th>df</th>
<th>Sig.</th>
<th>Exp(B)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Conditions</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ambiguity-Range vs.</td>
<td>-.334</td>
<td>.183</td>
<td>3.313</td>
<td>1</td>
<td>.069</td>
<td>.716</td>
</tr>
<tr>
<td>Risk</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Certain vs. Risk</td>
<td>-.883</td>
<td>.208</td>
<td>17.972</td>
<td>1</td>
<td>&lt;.0001</td>
<td>.414</td>
</tr>
<tr>
<td>Ambiguity-High vs.</td>
<td>1.202</td>
<td>.154</td>
<td>60.992</td>
<td>1</td>
<td>&lt;.0001</td>
<td>3.325</td>
</tr>
<tr>
<td>Risk</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ambiguity-Low vs.</td>
<td>.743</td>
<td>.158</td>
<td>22.215</td>
<td>1</td>
<td>&lt;.0001</td>
<td>2.103</td>
</tr>
<tr>
<td>Risk</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drop location</td>
<td></td>
<td></td>
<td>62.537</td>
<td>5</td>
<td>.000</td>
<td></td>
</tr>
<tr>
<td>Right vs. Lower right</td>
<td>-.272</td>
<td>.165</td>
<td>2.697</td>
<td>1</td>
<td>.101</td>
<td>.762</td>
</tr>
<tr>
<td>Upper left vs. Lower right</td>
<td>-.903</td>
<td>.177</td>
<td>26.021</td>
<td>1</td>
<td>.000</td>
<td>.406</td>
</tr>
<tr>
<td>Lower left vs. Lower right</td>
<td>-.257</td>
<td>.163</td>
<td>2.493</td>
<td>1</td>
<td>.114</td>
<td>.773</td>
</tr>
<tr>
<td>Left vs. Lower right</td>
<td>-1.319</td>
<td>.199</td>
<td>44.133</td>
<td>1</td>
<td>.000</td>
<td>.267</td>
</tr>
<tr>
<td>Upper right vs. Lower right</td>
<td>-.191</td>
<td>.164</td>
<td>1.357</td>
<td>1</td>
<td>.244</td>
<td>.826</td>
</tr>
<tr>
<td><strong>Monkey</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leo vs. Chris</td>
<td>.014</td>
<td>.156</td>
<td>.008</td>
<td>1</td>
<td>.928</td>
<td>1.014</td>
</tr>
<tr>
<td>Xenon vs. Chris</td>
<td>1.629</td>
<td>.131</td>
<td>154.156</td>
<td>1</td>
<td>.000</td>
<td>5.101</td>
</tr>
<tr>
<td>Constant</td>
<td>-2.402</td>
<td>.183</td>
<td>172.553</td>
<td>1</td>
<td>.000</td>
<td>.091</td>
</tr>
</tbody>
</table>
Appendix Table 7. Logistic regression for Chris.

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>S.E.</th>
<th>Wald</th>
<th>df</th>
<th>Sig.</th>
<th>Exp(B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drop location</td>
<td></td>
<td></td>
<td>24.616</td>
<td>5</td>
<td>.000</td>
<td></td>
</tr>
<tr>
<td>Right vs. Lower right</td>
<td>-.640</td>
<td>.471</td>
<td>1.849</td>
<td>1</td>
<td>.174</td>
<td>.527</td>
</tr>
<tr>
<td>Upper left vs. Lower right</td>
<td>.064</td>
<td>.381</td>
<td>.028</td>
<td>1</td>
<td>.866</td>
<td>1.066</td>
</tr>
<tr>
<td>Lower left vs. Lower right</td>
<td>.660</td>
<td>.354</td>
<td>3.476</td>
<td>1</td>
<td>.062</td>
<td>1.934</td>
</tr>
<tr>
<td>Left vs. Lower right</td>
<td>-.799</td>
<td>.450</td>
<td>3.152</td>
<td>1</td>
<td>.076</td>
<td>.450</td>
</tr>
<tr>
<td>Upper right vs. Lower right</td>
<td>.780</td>
<td>.342</td>
<td>5.216</td>
<td>1</td>
<td>.022</td>
<td>2.181</td>
</tr>
<tr>
<td>Conditions</td>
<td></td>
<td></td>
<td>9.806</td>
<td>4</td>
<td>.044</td>
<td></td>
</tr>
<tr>
<td>Ambiguity-Range vs. Risk</td>
<td>.579</td>
<td>.367</td>
<td>2.497</td>
<td>1</td>
<td>.114</td>
<td>1.785</td>
</tr>
<tr>
<td>Certain vs. Risk</td>
<td>-.128</td>
<td>.415</td>
<td>.096</td>
<td>1</td>
<td>.757</td>
<td>.879</td>
</tr>
<tr>
<td>Ambiguity-High vs. Risk</td>
<td>.822</td>
<td>.355</td>
<td>5.348</td>
<td>1</td>
<td>.021</td>
<td>2.274</td>
</tr>
<tr>
<td>Ambiguity-Low vs. Risk</td>
<td>.530</td>
<td>.373</td>
<td>2.020</td>
<td>1</td>
<td>.155</td>
<td>1.698</td>
</tr>
<tr>
<td>Constant</td>
<td>-2.994</td>
<td>.379</td>
<td>62.563</td>
<td>1</td>
<td>.000</td>
<td>.050</td>
</tr>
</tbody>
</table>
Appendix Table 8. Logistic regression for Leo.

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>S.E.</th>
<th>Wald</th>
<th>df</th>
<th>Sig.</th>
<th>Exp(B)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Drop location</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Right vs. Lower right</td>
<td>-2.723</td>
<td>.494</td>
<td>30.414</td>
<td>1</td>
<td>.000</td>
<td>.066</td>
</tr>
<tr>
<td>Upper left vs. Lower right</td>
<td>-3.896</td>
<td>.736</td>
<td>27.983</td>
<td>1</td>
<td>.000</td>
<td>.020</td>
</tr>
<tr>
<td>Lower left vs. Lower right</td>
<td>-.575</td>
<td>.262</td>
<td>4.809</td>
<td>1</td>
<td>.028</td>
<td>.563</td>
</tr>
<tr>
<td>Left vs. Lower right</td>
<td>-4.433</td>
<td>1.021</td>
<td>18.848</td>
<td>1</td>
<td>.000</td>
<td>.012</td>
</tr>
<tr>
<td>Upper right vs. Lower right</td>
<td>-20.173</td>
<td>2851.779</td>
<td>.000</td>
<td>1</td>
<td>.994</td>
<td>.000</td>
</tr>
<tr>
<td><strong>Conditions</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ambiguity-Range vs. Risk</td>
<td>.451</td>
<td>.440</td>
<td>1.050</td>
<td>1</td>
<td>.305</td>
<td>1.569</td>
</tr>
<tr>
<td>Certain vs. Risk</td>
<td>-1.810</td>
<td>.791</td>
<td>5.239</td>
<td>1</td>
<td>.022</td>
<td>.164</td>
</tr>
<tr>
<td>Ambiguity-High vs. Risk</td>
<td>1.382</td>
<td>.408</td>
<td>11.457</td>
<td>1</td>
<td>.001</td>
<td>3.984</td>
</tr>
<tr>
<td>Ambiguity-Low vs. Risk</td>
<td>1.714</td>
<td>.394</td>
<td>18.888</td>
<td>1</td>
<td>.000</td>
<td>5.553</td>
</tr>
<tr>
<td>Constant</td>
<td>-1.754</td>
<td>.355</td>
<td>24.377</td>
<td>1</td>
<td>.000</td>
<td>.173</td>
</tr>
</tbody>
</table>
Appendix Table 9. Logistic regression for Xenon.

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>S.E.</th>
<th>Wald</th>
<th>df</th>
<th>Sig.</th>
<th>Exp(B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drop location</td>
<td></td>
<td></td>
<td>54.303</td>
<td></td>
<td>.000</td>
<td></td>
</tr>
<tr>
<td>Right vs. Lower right</td>
<td>.762</td>
<td>.236</td>
<td>10.404</td>
<td>1</td>
<td>.001</td>
<td>2.144</td>
</tr>
<tr>
<td>Upper left vs. Lower right</td>
<td>-.334</td>
<td>.244</td>
<td>1.876</td>
<td>1</td>
<td>.171</td>
<td>.716</td>
</tr>
<tr>
<td>Lower left vs. Lower right</td>
<td>-.609</td>
<td>.256</td>
<td>5.663</td>
<td>1</td>
<td>.017</td>
<td>.544</td>
</tr>
<tr>
<td>Left vs. Lower right</td>
<td>-.714</td>
<td>.266</td>
<td>7.212</td>
<td>1</td>
<td>.007</td>
<td>.490</td>
</tr>
<tr>
<td>Upper right vs. Lower right</td>
<td>.498</td>
<td>.236</td>
<td>4.433</td>
<td>1</td>
<td>.035</td>
<td>1.645</td>
</tr>
<tr>
<td>Conditions</td>
<td></td>
<td></td>
<td>162.857</td>
<td></td>
<td>.000</td>
<td></td>
</tr>
<tr>
<td>Ambiguity-Range vs. Risk</td>
<td>-1.058</td>
<td>.261</td>
<td>16.446</td>
<td>1</td>
<td>.000</td>
<td>.347</td>
</tr>
<tr>
<td>Certain vs. Risk</td>
<td>-1.127</td>
<td>.265</td>
<td>18.064</td>
<td>1</td>
<td>.000</td>
<td>.324</td>
</tr>
<tr>
<td>Ambiguity-High vs. Risk</td>
<td>1.500</td>
<td>.207</td>
<td>52.696</td>
<td>1</td>
<td>.000</td>
<td>4.479</td>
</tr>
<tr>
<td>Ambiguity-Low vs. Risk</td>
<td>.503</td>
<td>.207</td>
<td>5.904</td>
<td>1</td>
<td>.015</td>
<td>1.654</td>
</tr>
<tr>
<td>Constant</td>
<td>-1.085</td>
<td>.214</td>
<td>25.650</td>
<td>1</td>
<td>.000</td>
<td>.338</td>
</tr>
</tbody>
</table>
Appendix Table 10. Number of correct trials out of non-dropped trial in Phase 4 and 5

<table>
<thead>
<tr>
<th>Conditions</th>
<th>Correct/non-drop</th>
<th>Total correct/non-drop</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Leo</td>
<td>Chris</td>
</tr>
<tr>
<td>Risk</td>
<td>119/230</td>
<td>101/227</td>
</tr>
<tr>
<td>Ambiguity-Range</td>
<td>110/225</td>
<td>119/218</td>
</tr>
<tr>
<td>Ambiguity-High</td>
<td>104/212</td>
<td>110/213</td>
</tr>
<tr>
<td>Ambiguity-Low</td>
<td>99/200</td>
<td>113/220</td>
</tr>
<tr>
<td>Certain</td>
<td>232/238</td>
<td>207/228</td>
</tr>
<tr>
<td>Risk(Phase 5)</td>
<td>116/226</td>
<td>N.A.</td>
</tr>
</tbody>
</table>

Appendix 4 Pan-Fragaszy softmax model explains ambiguity aversion in the decline paradigm and urn paradigm.

Now we will show how the equation (4) can be applied to explain the ambiguity aversion in the two experimental paradigms. In the first paradigm with $3 fixed payoff for dropping, the probability of dropping a trial in the risky uncertainty condition is:

\[ P_{\text{drop}(\text{risk})} = \frac{e^{-\beta/(1\times3)}}{e^{-\beta/(0.5\times10)} + e^{-\beta/(0.5\times10)} + e^{-\beta/(1\times3)}} = \frac{e^{-\beta/(1\times3)}}{2e^{-\beta/(0.5\times10)} + e^{-\beta/(1\times3)}}. \]  

Assume that the subjective probability of a card being red or blue follows a discrete uniform distribution within the range of \([0, 1]\) at \(m+1\) locations (where \(m\) is a positive integer). That is, subjective probability is a second-order probability distribution of its possible values (Marschak, 1975, cited in Camerer and Weber (1992 p. 327). For all \(m-1\) judged probabilities of one choice within the range of \((0,1)\), the \(i^{th}\) judged probability is \(i/m\) and each \(i^{th}\) judged subjective
probability has a weight of $1/(m-1)$ in the probability of choosing this choice. The probability of choosing to bet that a card is red (or blue) is:

$$P_{\text{bet red (ambiguous)}} \propto \sum_{i=1}^{m-1} \frac{1}{m-1} \times e^{-\beta/(\frac{1}{m} \times 10)} = \frac{1}{m-1} \sum_{i=1}^{m-1} e^{-\beta/(\frac{1}{m} \times 10)}. \quad (S.2)$$

Thus the probability of dropping a trial in the ambiguous uncertainty condition is:

$$P_{\text{drop (ambiguous)}} = \frac{\frac{\beta}{1 \times 3}}{\frac{1}{m-1} \sum_{i=1}^{m-1} e^{-\beta/(\frac{1}{m} \times 10)} + \frac{1}{m-1} \sum_{i=1}^{m-1} e^{-\beta/(\frac{1}{m} \times 10)} + e^{-\frac{\beta}{1 \times 3}}}$$

$$= \frac{e^{-\beta/(1 \times 3)}}{\frac{1}{m-1} \sum_{i=1}^{m-1} e^{-\beta/(\frac{1}{m} \times 10)} + e^{-\beta/(1 \times 3)}}. \quad (S.3)$$

To show that $P_{\text{drop (risk)}} \leq P_{\text{drop (ambiguous)}}$ is equivalent to show that:

$$\frac{e^{-\beta/(1 \times 3)}}{2e^{\beta/(0.5 \times 10)} + e^{\beta/(1 \times 3)}} < \frac{2e^{-\beta/(1 \times 3)}}{\frac{1}{m-1} \sum_{i=1}^{m-1} e^{-\beta/(\frac{1}{m} \times 10)} + e^{-\beta/(1 \times 3)}}$$

$$2e^{-\beta/(0.5 \times 10)} > \frac{2}{m-1} \sum_{i=1}^{m-1} e^{-\beta/(\frac{1}{m} \times 10)}$$

$$(m - 1)e^{-\beta/(0.5 \times 10)} > \sum_{i=1}^{m-1} e^{-\beta/(\frac{1}{m} \times 10)} \quad (S.4)$$

The following special case shows that for a possible chosen fixed value of $\beta$ and for $m$ up to 10000, inequality (S.4) holds. Let $\beta= 1$, (S.4) becomes:

$$(m - 1)e^{-1/0.5 \times 10} > \sum_{i=1}^{m-1} e^{-\frac{1}{m} \times 10}. \quad (S.5)$$

When $m = 3$, that is, when the judged probability of a card being red (or blue) is equally likely to be 0, 1/3, 2/3, or 1, inequality (S.5) holds:

$$2e^{-1/(0.5 \times 10)} = 1.6374 > \sum_{i=1}^{m-1} e^{-\frac{1}{m}} = e^{-1/(\frac{1}{3} \times 10)} + e^{-1/(\frac{2}{3} \times 10)} = 1.6015.$$
Similarly, when \( m = 4 \),

\[
3e^{-1/(5 \times 10)} = 2.456 > e^{-1/(\frac{1}{4} \times 10)} + e^{-1/(\frac{2}{4} \times 10)} + e^{-1/(\frac{3}{4} \times 10)} = 2.36.
\]

...

When \( m = 10,001 \),

\[
10000e^{-0.5} = 6053.3 > \sum_{i=1}^{10000} e^{-\frac{i}{10000 \times 10}} = 999.4547.
\]

In the second paradigm, without a fixed payoff for dropping, there are actually four choices a subject can choose from the set: to bet on red from the risky deck, to bet on blue from the risky deck, to bet on red from the ambiguous deck, and to bet on blue from the ambiguous deck. According to the third axiom of the three axioms of probability (Feller, 1957), the probability of betting from the ambiguous deck = the probability of betting on red from the ambiguous deck + the probability of betting blue from the ambiguous deck. Similarly, the probability of betting from the risky deck = the probability of betting on red from the risky deck + the probability of betting blue from the risky deck:

\[
P(\text{risk}) = P(\text{bet red (risk)}) + P(\text{bet blue (risk)})
\]

Or

\[
P(\text{risk}) = \frac{e^{-\beta/(0.5 \times 10)} + e^{-\beta/(1.5 \times 10)}}{e^{-\beta/(0.5 \times 10)} + e^{-\beta/(1.5 \times 10)} + \frac{1}{m-1} \sum_{i=1}^{m-1} e^{-\beta/(\frac{i}{m} \times 10)} + \frac{1}{m-1} \sum_{i=1}^{m-1} e^{-\beta/(\frac{m-i}{m} \times 10)}}.
\]

Similarly,

\[
P(\text{ambi}) = \frac{\frac{1}{m-1} \sum_{i=1}^{m-1} e^{-\beta/(\frac{i}{m} \times 10)} + \frac{1}{m-1} \sum_{i=1}^{m-1} e^{-\beta/(\frac{m-i}{m} \times 10)}}{e^{-\beta/(0.5 \times 10)} + e^{-\beta/(1.5 \times 10)} + \frac{1}{m-1} \sum_{i=1}^{m-1} e^{-\beta/(\frac{i}{m} \times 10)} + \frac{1}{m-1} \sum_{i=1}^{m-1} e^{-\beta/(\frac{m-i}{m} \times 10)}}.
\]

To show that \( P(\text{risk}) > P(\text{ambi}) \), is equivalent to show the inequality (S.4):

\[
(m - 1)e^{-\beta/(0.5 \times 10)} > \sum_{i=1}^{m-1} e^{-\beta/(\frac{i}{m} \times 10)},
\]

which was shown in the first paradigm.
Since we use integration, whether the edge of the judged probability includes zero or one does not matter anymore. Analogous to inequality (S.4), to show ambiguity aversion is to show inequality (S.6) in a special case $\beta = 1$:

$$e^{-\beta/(0.5\times10)} > \int_0^1 e^{-\frac{\beta}{p\times10}}d(p)$$  \hspace{1cm} (S.6)

$$e^{-\beta/(0.5\times10)} = 0.8187 > \int_0^1 e^{-\frac{\beta}{p\times10}}d(p) = 0.722545$$

which holds.

Appendix 5 Challenges of Kahn and Sarin model

Kahn and Sarin (1988)'s decision weight model (KS model, hereafter) also weights second order subjective probability nonlinearly. Their second order subjective probability is also continuously uniform over 0 to 1, i.e., p.d.f. $\Phi(p) = 1$.

Classical subjective expected utility (SEU, hereafter) defines an ambiguous event $E$ in the lottery $L$ with probability $p$ and money value $x$ as

$$SEU(L) = \int_{p=0}^{1} p\Phi(p)d\mu(x) = \tilde{p}\mu(x) \hspace{1cm} (S.7)$$

in which $u$ is the value function of money $x$.

KS model define the value ($V$) of an ambiguous event $E$ in the lottery $L$ as

$$V(L) = w(E)u(x) \hspace{1cm} (S.8),$$

in which $w(E)$ is called decision weight function in KS model. Comparison of KS model (equation S.8) with classical SEU (S.7) will reveal that the value function in KS model, $V(L)$, is analogous to the subjective expected utility, $SEU(L)$, and $u(x)$ in KS model is the value function $u(x)$ in the SEU. Please note that value function in KS model, $V(L)$, is actually a function of utility which deals not only with money value $x$, but also with probability $p$. Note that in
prospect theory, the value function only refers to psychological perception of money value, i.e., 
\( u(x) \), not the product of the perception of money value and perception of probability.

If one compares classical SEU with prospect theory, \( u(x) \) is the value function in both 
classical SEU and prospect theory. In S.8, \( u(x) \) deals only with money value \( x \), and can be 
perceived either the way SEU proposes or Prospect theory proposes. \( u(x) \) is independent of \( p \), so 
\( u(x) \) can be treated as a constant in the integration. We focus on how the probability part is 
perceived. In prospect theory, the perception of pure risk probability “p” is always weighted, 
denoted by \( w(p) \). If \( p = w(p) \), then the subjective perception of probability matches the objective 
probability, which is not the case in prospect theory. In prospect theory, \( p \) in classical SEU will 
replaced by \( w(p) \) in prospect theory. If one rewrites (S.7) with the concept of probability 
weighting of pure risk, it will be:

\[
\text{SEU}(L) = \int_{p=0}^{1} w(p) \Phi(p) d\mu(x) = \int_{p=0}^{1} w(p) d\mu(x) = \{\int_{p=0}^{1} w(p) dp\} \ast u(x) \quad (S.9)
\]

If one compares S.9 from SEU to S.8 from KS model, one will find the decision weight of event 
\( E \) in KS model, \( w(E) \), is cumulative value of the weighting function of probability in prospect 
theory:

\[
w(E) = \int_{p=0}^{1} w(p) dp \quad (S.10)
\]

In KS model, the decision weight function, \( w(E) \), is:

\[
w(E) = \bar{p} + \int_{p=0}^{1} (p - \bar{p}) e^{\frac{-\lambda(p-\bar{p})}{\sigma}} \Phi(p) dp
\]

\[
w(E) = \int_{p=0}^{1} p \Phi(p) dp + \int_{p=0}^{1} (p - \bar{p}) e^{\frac{-\lambda(p-\bar{p})}{\sigma}} \Phi(p) dp
\]

According to the sum property of integral:

\[
w(E) = \int_{p=0}^{1} \{ p + (p - \bar{p}) e^{\frac{-\lambda(p-\bar{p})}{\sigma}} \} \Phi(p) dp
\]
Remember that p.d.f. \( \Phi(p) = 1 \),

\[
w(E) = \int_{p=0}^{1} \{ p + (p - \bar{p}) e^{\frac{-\lambda(p - \bar{p})}{\sigma}} \} dp \quad (S.11)
\]

Comparing S.11 from the KS model to S.10 from prospect theory, the integrand in S.11 is exactly in the place of probability weighting function of prospect theory.

\[
w(p) = p + (p - \bar{p}) e^{\frac{-\lambda(p - \bar{p})}{\sigma}} \quad (S.12)
\]

where \( \lambda \) reflects an individual’s attitude toward ambiguity in a given context in which \( \lambda > 0 \) reflects ambiguity aversion and \( \sigma \) is the standard deviation of the second order probability distribution \( \Phi(p) \) in KS model.

Given that \( \bar{p} = 0.5 = \int_{p=0}^{1} p\Phi(p) dp \), the probability density function of the second order probability distribution \( \Phi(p) = 1 \) (i.e., the subject judged probability spread evenly from 0 to 1, a continuous uniform distribution), \( \sigma = \sqrt{1/12} \) and \( \lambda = 1 \), the weighting function of probability in Kahn-Sarin (1988)’s model is:

\[
w(E) = \int_{p=0}^{1} p dp + \int_{p=0}^{1} (p - 0.5) e^{\sqrt{T}(0.5-p)} dp
\]

According to sum property of integrals:

\[
w(E) = \int_{p=0}^{1} [p + (p - 0.5) e^{\sqrt{T}(0.5-p)}] dp
\]

\[= 0.11491 < 0.5, \text{ ambiguity averse.}\]

In this example the \( w(p) = p + (p - 0.5) e^{\sqrt{T}(0.5-p)} \) is depicted in Appendix Figure 6. This weight function of probability according to KS model (1) takes negative values when \( p \) is small, and (2) is convex upward. Those two properties are in conflict with prospect theory. In prospect theory, the probability weighting function does not take negative values and is concave upward. Proponents of the KS model may argue that their weighting function \( w(E) \) deals with attitude
toward ambiguity and is different from prospect theory. How any kind of weight of attitude toward the pure risk part of probability will take negative values remains a question for proponents of KS model to answer.

Appendix Figure 6. Probability weighting function in prospect theory derived from in KS model.
References

Journal of Political Economy, 72, 62-73.


Econometrica, 73(6), 1849-1892.


---

1 formerly known as *Cebus apella*

1 B value is the logistic coefficient similar to the b value in linear regression.