Probability Weighting in Decision Under Ambiguity

Thiago Cyfer Goularte & Rodrigo De Losso da Silveira Bueno

April 15, 2023

Abstract

We propose second-order belief model of decision-making under ambiguity, that extends the Salience Model of choice under risk (Bordalo et al., 2012). Our model is able to predict stylized facts of the literature, such as likelihood insensitivity, the fact that people tend to overestimate(underestimate) the expected probability of low(high)-likelihood events to happen. We also predict the fourfold pattern of ambiguity attitudes, where decision-makers are usually ambiguity averse for bets on high probability gains and low probability losses, and ambiguity seeking otherwise. A key feature of our model is that ambiguity attitude is a result of the combination of: (i) the outcome domain (gain or loss) of a bet; (ii) the expected probability of the events that are relevant to the outcome of the bet; (iii) how many states of the world are deemed possible by the Decision-Maker. So, there is no need for additional assumptions on our model to predict that ambiguity attitudes change according to those three factors, in agreement with the experimental and empirical literature. Furthermore, we include the context - represented by characteristics of the choice set of the decision-making problem - as a variable that may affect preferences, in the same spirit of the original Salience Theory.

1 Introduction

Since Ellsberg's (1961) seminal thought experiments for choice under ambiguity, the decision-making behavior under circumstances of high uncertainty about the probabilities of outcomes or relevant events for a choice have been studied. Specifically, theoretical models that deal with decision under ambiguity try to accomodate and/or describe decision-makers (DMs) behavior when dealing with decisions involving uncertainty where assessing a unique probability distribution (even if subjective) of events relevant to the outcomes of the possible courses of action is particularly complex, or information about such probabilities is scarce and/or imprecise. The presence of ambiguity may affect preferences in a non-obvious way (Machina, 2009). Moreover, ambiguity attitudes have received much attention as a possible explanation for economic phenomena such as the equity premium puzzle (Rieger & Wang, 2012), stock market non-participation (Antoniou et al., 2015; Dimmock et al., 2016) and home bias in portfolio choices (Dimmock et al., 2016; Ardalan, 2019).

We contribute to the theoretical literature by axiomatizing properties of a probability weighting function that reflects DMs behavior for second-order probability represenations of decision under ambiguity problems. The probability weighting function that we propose, together with common value/utility functions such as Prospect-Theory (Kahneman & Tversky, 1979) is able to predict many empirical regularities of the experimental literature, such as: (i) likelihood insensitivity, the idea that under ambiguity DMs tend to distort the weighting of events in the evaluation of an act in the direction of a naive probability distribution (for example, a 50-50 naive probability distribution for a bet with two possible outcomes) (Dimmock et al., 2013; Trautmann & van de Kuilen, 2017); (ii) a fourfold pattern of ambiguity attitude, that describes how the ambiguity attitudes of individuals varies as a function of the expected probability of an event and the domain of the outcome (gains or losses) associated with the event occurrence for a given act. For example, when comparing a risky (with objectively know probability of win) and an ambiguous bet (with imprecise information on the probability of win) that involve the possibility of gaining a positive value with small expected probability, DMs usually tend to choose the ambiguous act (be ambiguity seeking). However, if the bet now involves the same gain, but events with mid to high expected probability of the win outcome, now the DM will usually prefer the risky bet (Trautmann & van de Kuilen, 2015).

More specifically, our model can be interpreted as relating the stylized facts above in the following way: once faced with imprecise information about the probabilities of events that are relevant to the outcomes of a number of alternative courses of action, the decision-maker "fills" the information gap with other information that she is able to assess from the choice set, such as the number of possible events. This use of the number of possible states of the world induced by the choice set description causes likelihood insensitivity, interpreted as the DMs distortion in weighting of the events toward a naive equal probability for each relevant event deemed possible. Therefore, likelihood insensitivity causes the fourfold pattern of ambiguity attitudes: for ambiguous acts, it causes an underweight of high likelihood events, resulting in a pessimistic (optimistic) view of expected gains (losses), causing ambiguity aversion (seeking) behavior. Conversely, ambiguous acts contingent on low likelihood events are overweighted, causing ambiguity seeking (averse) behavior for gains (losses).

Thus, ambiguity attitude need not to be an assumption made before applying the model, as in other second-order belief models (e.g., Klibanoff et al., 2005). In our model, the ambiguity attitude is a result of the combination of the outcome domain, expected probability of events, and how many states of the world are deemed possible (i.e., are non-null) by the decision-maker.

Moreover, our probability weighting function is also able to account for puzzles related to the nonseparability of the ambiguity degree of combined events, such as Machina (2009) reversals. Machina reversal examples highlight that, depending on the outcome-event pairs associated with each bet, it is possible that two bets based on the same states of the world to have different degrees of ambiguity (or even one of them being unambiguous and the other one ambiguous), and that is specifically related to how the acts in a choice set associated events and outcomes.

The central idea of our paper of the choice set outcomes shaping the DM's perception of the state-space has been applied to other decision contexts. Specifically, the psychological rationale of bottom-up attention, i.e., a stimulus caused by the specific context of a choice problem attracting the decision maker's attention "bottom up," automatically and involuntarily, is the main driver of Salience Theory of Choice Under Risk (STR) (Bordalo et al., 2021). As may be expected, our probability weighting function is then particularly well suited to be applied to extend STR for Choice Under Ambiguity. We explore this way to apply the probability weighting function as a theory of Choice Under Ambiguity that has STR as a special case - when the decision collapses to risky acts. Salience Theory has also already been tested and used to explain a variety of stylized facts, such as the tendency to take right-skewed risks and avoid left-skewed ones (Kahneman & Tversky, 1979) and Allais paradoxes (Bordalo et al., 2012).

To the best of our knowledge, our model is the first one that is able to predict likelihood insensitivity as a function of characteristics of the choice set. Moreover, our model is able to simultaneously deal with Machina reversals (Gul & Pesendorfer, 2013), and not only to accommodate (Klibanoff et al., 2005; Chateneuf & Faro, 2009; Chateneauf, Eichberger & Grant (2007); Wakker, 2010) but predict the fourfold pattern of ambiguity attitudes observed in the empirical literature. We take advantage of the rapid growth in the experimental and empirical literature on decision-making under ambiguity to make our model more in line with the regularities in DMs behavior found in the literature.

The rest of this paper is organized as follows. Section 2 presents a running Ellsberg-urn example that is going to be used throughout the paper to explain the concepts and consequences of our model; Section 3 gives the preliminaries of our model: how our second-order belief approach describes decision under ambiguity and some background on Salience Theory. Section 4 presents our Context-Weighting model, already applying it as a generalization of Salience Theory of Choice Under Risk (Bordalo et al, 2012). In the subsections within this section, we give the model's preliminaries, postulates about how DMs interpret second-order probability distributions and, whenever possible, illustrate how theses postulates affect DMs preferences and weighting of act outcomes. For concreteness, we also provide a parametric example of a Context-weighting function, and apply it to our running example and some chosen modifications of it that help to illustrate our model's properties. Section 5 illustrates how Choice set characteristics, DMs beliefs (i.e., second order characteristics) and the Outcome Domain of act payoffs influence ambiguity attitude in our model. We also analyze the Machina reversal problem and how events with correlated probability are dealt with in our model. We follow with Section 6, where we compare our model's characteristics and results with other popular second-order belief models in the literature. Finally, in Section 7 we draw conclusions and give suggestions for future research.

2 Running Example

In this section we present an Ellsberg urn like example to illustrate how the model can be applied Suppose there is an urn with 12 balls. The decision-maker (DM) knows for sure that there are 3 red balls and 3 green balls in the urn. The 6 remaining balls can be either black or yellow, in proportion unknown to the DM. The DM interprets her lack of information about the right proportion of black and yellow balls contained in the urn as each possible combination being equally likely. For example, the DM believes the probability of the urn having 6 yellow balls (and no black balls) is the same as the probability of it having 2 yellow balls (and 4 black balls). The DM is then offered the bets f_0 and f_α, contingent on the color of a single ball drawn from the urn. Table 1 below describes the monetary outcomes associated with each possible color of the drawn ball:

	Red ($p = 3/12$)	Yellow ($p \in [0, 6/12]$)	Black ($p \in [0, 6/12]$)	Green ($p = 3/12$)
f_0	100	0	0	150
f_α	0	100	0	150

Table 1: An Ellsberg-like ambiguity example (Example 1). We also refer to the example of this table as the running example, throughout the text.

It is easy to see that the events "a red ball is drawn" (E_r) and "a green ball is drawn" (E_g) are unambiguous, whilst "a yellow ball is drawn" (E_y) and "a black ball is drawn" (E_b) are ambiguous. However, note that the outcomes of act f_0 are only contingent on unambiguous events: (i) E_r (wins \$100 with probability 3/12); E_g (wins \$150 with probability 3/12); $E_{(y \vee b)}$ (gets \$0 with probability 6/12). On the other hand, some of f_α outcomes are contingent on ambiguous events. If the DM chooses f_α, she can: win \$100 with some unknown probability, that can between 0 and 6/12 (E_y obtains); get \$0 with some unknown probability, that can between 3/12 and 9/12 ($E_{(r \vee y)}$ obtains); win \$150 with probability 3/12 (E_g obtains) - this last one the only unambiguous event-contingent outcome. Slightly abusing terminology, we will refer to f_0 as an unambiguous act, and f_α as an ambiguous act, when referring to our running example.

3 Preliminaries

3.1 Decision Under Ambiguity Representation

A choice under ambiguity is described by:

(i) a non-empty (countable) finite set S representing the states of the world, along with its Borel σ-algebra $\mathcal{A}(S)$;

(ii) An arbitrary set of acts $F = \{f_1, ..., f_I\}$[1], which constitutes the decision-maker's (DM) choice set. Each act is a mapping $f : S \to X$, that assigns consequences $x \in X$ to states of the world $s \in S$;

(iii) An arbitrary finite set of consequences $X \subset \mathbb{R}$, representing the subset of payoffs induced by the choice set F[2];

(iv) The set $\Delta(S)$ of all probability distribution measures on the measurable space $(S, \mathcal{A}(S))$, along with its Borel σ-algebra $\mathcal{A}(\Delta(S))$;

(v) A probability distribution measure $\mu : \mathcal{A}(\Delta(S)) \to \mathbb{P}$ defined on the measurable space $(\Delta(S), \mathcal{A}(\Delta(S)))$. This measure assigns a second-order probability to each first-order probability distribution $P \in \Delta(S)$.

Events E are subsets of S. For conciseness, we sometimes denote by xEy any act that yields x if event E obtains, and y otherwise. We define a null event $A \in \mathcal{A}(S)$ as one satisfying the condition $f_iAg \sim f_jAg$, for all $f_i, f_j, g \in F$. This definition is similar to other Decision Under Ambiguity models, such as Gul and Pesendorfer (2013). Similarly, a null probability distribution is any $\pi_0 \in \Delta(S)$ for which $\mu(\pi_0) = 0$. We define $\Pi = \{\pi \in \Delta(S) : \mu(\pi) > 0\}$ as the set of non-null probability distributions and $\mathcal{A}(\Pi)$ as its Borel σ-algebra. We typically assume that Π contains N elements, i.e., $\Pi = \{\pi_1, ..., \pi_N\}$. Slightly abusing terminology, we also loosely refer to non-null probability distributions as plausible distributions throughout the text. This setting follows closely other second-order beliefs models, such as Klibanoff et al. (2005).

We denote the marginal compound probability of an event E under distributions $\mu(\Pi)$ and $\pi_i(E)$ as $\mu(\pi_i(E)) = \mu(\pi_i) \cdot \pi_i(E)$. The expected value of the probability of event E is represented as $\mathbb{E}\mu(\Pi(E)) = \sum_{i=1}^{N} \mu(\pi_i) \cdot \pi_i(E)$, and its standard deviation is $\sigma(\mu(\Pi(E))) = \sqrt{\sum_{i=1}^{N} \mu(\pi_i)(\pi_i(E) - \mathbb{E}_\mu(\Pi(E)))^2}$. These concepts are frequently employed throughout the paper, as the valuation of an act contingent on ambiguous events is dependent on the second-order μ distribution representation of the DM's beliefs.

We now introduce the concept of an Act-Induced Partition. This definition is especially helpful when dealing with sets of acts that have outcomes contingent on distinct events, since ambiguity is significant only for events that play a role in determining an act's outcome. In our ongoing example, this distinction is easily made: for instance, event E_y is important for determining the outcome of act f_α, but only $E_{(y \vee b)}$ (and not the events E_y and E_b separately) is essential for determining the outcomes of f_0. The Act-induced partition definition is particularly useful for handling such cases. From now on, we denote the partition induced by an act f_i as $\mathcal{P}^i$.

[1] To avoid technical difficulties that do not add much value to our model, we assume: (a) that DM's understand acts that yield the same outcomes contingent on each state $s \in S$ as reducible to a choice set without act-duplicates, i.e., $F = \{f_i \in F' : \forall f_j \in F' f_i \neq f_j\}$; (b) the choice set is non-trivial, i.e., $f_i \neq f_j$ for some $f_i, f_j \in F$

[2] formally, $X = \bigcup_{i=1}^{I} f_i(S)$, i.e., is the union of the images of the available acts $f_i \in F$ for all possible states of the world $s \in S$.

Definition 1 (Act-Induced Partition) Let $f_i : S \to X$ be an act such that $f_i \in F$. Then, for any $x \in X$, define $E^i_x = \{s \in S : f_i(E) = x\}$ as the set of events in S that are mapped to $x \in X$ by f_i. Then, the f_i Act-Induced Partition $\mathcal{P}^i$ is the partition of S that contains every E^i_x for each $x \in X$, i.e., $\mathcal{P}^i = \{E^i_x : x \in X\}$.

In other words, $\mathcal{P}^i$ is a partition of the state space S such that each subset of the partition corresponds to a unique outcome in X under the act f_i. This partition represents the way in which f_i maps the events of the state space to the outcomes in X.

Note that, by the definition of a partition of a set, and assuming the choice set F is non-empty, then every $E^i_x \in \mathcal{P}^i$ is non-null. Second, for any act contingent on an ambiguous event, there must be at least two events in its' Act-Induced Partition (otherwise, the DM would know that the only event $s \in E^i_x$ obtains with probability 1). That is, for a discrete state-space S the cardinality of $\mathcal{P}^i$ ($|\mathcal{P}^i| \in \mathbb{N}$) is greater or equal to 2 for any ambiguous act $f_i \in F$.

For concreteness, let's apply the definition to example 1. The act-induced partitions for each act are determined by the outcomes induced by the acts, $\mathcal{P}^0 = \{\{E_r\}, \{E_y, E_b\}, \{E_g\}\}$ and $\mathcal{P}^\alpha = \{\{E_r, E_b\}, \{E_y\}, \{E_g\}\}$.

In agreement with models such as Prospect Theory (PT) and the Salience Theory of Choice Under Risk (STR) model, the DM uses a value function $v : X \to \mathbb{R}$ to evaluate payoffs, relative to the reference point of zero[3]. We concentrate on a PT-like (as in Kahneman & Tversky, 1979) value function exhibiting loss aversion, although alternative functions such as linear utility (with respect to a reference point) (Bordalo et al., 2012) can also be utilized with similar outcomes. Following STR, we denote x^i_s as the monetary payoff yielded by the act f_i if state $s \in S$ obtains. $\mathbf{x}_s = \{x^1_s, ..., x^I_s\}$ is the set of payoffs implied by each act in F, conditional on the state s being true. $\mathbf{x}^{-i}_s = \{x^j_s\}_{j \neq i}$ is the set of payoffs implied by each act in $(F \setminus f_i)$ - i.e., all acts in F except act f_i - whenever the true state is s.

According to our preliminaries, our running example presented in the previous section could then be represented in the following way:

(i) $S = \{E_r, E_y, E_b, E_g\}$ is the state-space. S Borel σ-algebra is
$\mathcal{A}(S) = \{\emptyset, \{E_r\}, \{E_y\}, \{E_b\}, \{E_g\}, \{E_r, E_y\}, \{E_r, E_b\}, \{E_r, E_g\}, \{E_b, E_y\}, \{E_b, E_g\}, \{E_y, E_g\}$
$, \{E_r, E_y, E_b\}, \{E_r, E_y, E_g\}, \{E_y, E_b, E_g\}, S\}$;

(ii) The choice set is $F = \{f_0, f_\alpha\}$;

(iii) $X = \{0, 100, 150\}$ is the set of consequences;

(iv) Π is the set of all probability functions associated with the measurable space $(S, \mathcal{A}(S))$, i.e., $\Delta(S) = \{\pi(r), \pi(y), \pi(b), \pi(g) \in [0, 1] : \pi(r) + \pi(y) + \pi(b) + \pi(g) = 1\}$, where $\pi(r), \pi(y), \pi(b), \pi(g)$ are the (first-order) probabilities of E_r, E_y, E_b, E_g being true, respectively. The set of non-null elements of $\Delta(S)$ is $\Pi = \{\Pi_1, ..., \Pi_7\}$, with elements given in the Table 2 below;

(v) Assuming that the DM believes that each plausible ball composition of the urn is equally likely to be true, $\mu(\Pi_1) = \mu(\Pi_2) = \mu(\Pi_3) = \mu(\Pi_4) = \mu(\Pi_5) = \mu(\Pi_6) = \mu(\Pi_7) = \frac{1}{7}$, and $\mu(\Pi_q) = 0$, for all other $\Pi_q \in \Delta(S)$ such that $\Pi_q \cap (\bigcup_{i=1}^{7} \Pi_i) = \emptyset$[4].

3.2 Salience Representation

Following the STR Model, we define a continuous and bounded salience function for each state of the world $s \in S$ as $\omega : F \to \mathbb{R}$. Given S and F, a function $\omega(x^i_s, x^{-i}_s)$ is considered a salience function if it

[3] Without loss of generality, other reference points may also be employed in our model.

[4] we will discuss second-order belief formation in section 5.2

	π_r	π_y	π_b	π_g
Π_1	3/12	0	6/12	3/12
Π_2	3/12	1/12	5/12	3/12
Π_3	3/12	2/12	4/12	3/12
Π_4	3/12	3/12	3/12	3/12
Π_5	3/12	4/12	2/12	3/12
Π_6	3/12	5/12	1/12	3/12
Π_7	3/12	6/12	0	3/12

Table 2: Non-null first-order probability distributions in our running example.

satisfies:

(i) Ordering: if $x_s^i = \max \mathbf{x}_s$, then for any $\epsilon, \epsilon' \geq 0$, with at least one strict inequality:

$\omega(\mathbf{x}_s^i + \epsilon, \mathbf{x}_s^{-i} - \epsilon') > \omega(\mathbf{x}_s^i, \mathbf{x}_s^{-i})$

if $x_s^i = \min \mathbf{x}_s$, then for any $\epsilon, \epsilon' \geq 0$, with at least one strict inequality:

$\omega(\mathbf{x}_s^i - \epsilon, \mathbf{x}_s^{-i} + \epsilon') > \omega(\mathbf{x}_s^i, \mathbf{x}_s^{-i})$

(ii) Diminishing sensitivity: if $x_s^j > 0$ for all j, then for any $\epsilon > 0$,

$\omega(\mathbf{x}_s^i + \epsilon, \mathbf{x}_s^{-i} + \epsilon') < \omega(x_s^i, x_s^{-i})$

(iii) Reflection: for any two states $s, s' \in S$ s.t. $x_s^j > 0$ for all j, we have

$\omega(\mathbf{x}_s^i, \mathbf{x}_s^{-i}) < \omega(\mathbf{x}_{s'}^i, \mathbf{x}_{s'}^{-i}) \Leftrightarrow \omega(-\mathbf{x}_s^i, -\mathbf{x}_s^{-i}) < \omega(-\mathbf{x}_{s'}^i, -\mathbf{x}_{s'}^{-i})$

(i) and (ii) are the key properties to explain anomalies such as the Allais Paradox, while (iii) plays a role in the determining the model's predictions when there are negative payoffs. Another important remark is that the average $\bar{x}_s^{-i} = \sum_{j \neq i} \frac{1}{I-1} x_s^j$ can substitute the set $\mathbf{x}_s^{-i}$ as an argument of the salience function, while keeping the properties (i), (ii) and (iii). In other words, $\omega(x_s^i, \bar{x}_s^{-i})$ is also a salience function (Bordalo et al., 2012). Without loss of generality, we will refer to $\omega(x_s^i, \bar{x}_s^{-i})$-type salience functions in our examples throughout the paper.

The salience function is meant to represent how a stimulus attracts the DM's attention "bottom up", that is, how perceived characteristics of a state of the world within a specific set of possible courses of action may impact decision-making. This concept contrasts with the traditional economic approach, which views attention as either unlimited or optimally allocated "top-down" based on current goals and expectations. This approach does not highlight that "bottom up" stimulus-driven attention may compete with the DM's "top down" goals (Bordalo et al., 2022). As Kahneman (2011, p. 324) puts it, "our mind has a useful capability to focus on whatever is odd, different or unusual". Salience Theory calls the payoffs that draw the decision maker's attention "salient".

An example is a DM confronted with the decision of using \$10 that he has in his pocket to buy a lottery ticket for a 0.001% chance of winning a \$1,000,000 dollar prize, or investing that money for a sure outcome of \$11 (the inicial \$10 plus a \$1 return). Even though the expected value of betting in the lottery is \$0 ($= 0.001\% \times (\$1,000,000 - \$10) + 99.999\% \times (\$-10)$), clearly less than the sure \$11 outcome of investing. In the standard Expected Utility Theory approach, it is evident that the risky bet's expected value being lower than the sure value of investing implies that any individual who is not risk-seeking should choose to invest. However, in reality, just the perspective of winning such a huge prize may tempt the DM to bet in the lottery instead.

That is an example of the high contrast between an outcome (receiving the lottery's prize), conditional on a state of the world being true (the numbers drawn in the prize draw match the DM's ticket) and a course of action chosen (buying the lottery ticket), compared to the same state outcome when a different course of action is taken (getting \$11 for investing). Salience theory interprets this as the state "the numbers drawn in the prize draw match the lottery ticket" being highly salient, because different choices lead to wildly different outcomes if this state is true. As a result, the state of the lottery prize going to the available ticket may have more weight in the DM's evaluation of his options, making her more inclined to buy the ticket for a chance of winning that huge prize, as opposed to the fear of regretting that she did not buy the lottery ticket that would have won the lottery.

A slight change to our running example illustrates how salience may affect weight the weight given by the DM to each state, conditional on the acts' outcomes. Say we add $z > 0$ to the outcome associated with the unambiguous state s_r for both acts (f_0, f_α) (Table 2 below).

	Red (E_r)	Yellow (E_y)	Black (E_b)	Green (E_g)
f_0	100+z	0	0	150
f_α	0+z	100	0	150

Table 3: An Ellsberg-like ambiguity example (Example 2). The difference between this example and our running example is that $z > 0$ is added to the outcomes associated with E_r being true, for both f_0 and f_α acts.

Then, by diminishing sensitivity, $\omega(x_{E_r}) < \omega(x_{E_y})$. That would result in an overweight of the difference in outcomes when E_y is true. That is, the difference between \$100 and \$0 outcomes in favor of f_α when a yellow ball is drawn seems now more attractive than the $\$100+z$ versus $\$0+z$ difference in outcomes when a red ball is drawn. That would lead to an overweight of the E_y event in the DM's choice, resulting in a more favorable view of the f_α option. That is, if in example 1 $f_0 \succ f_\alpha$, then for Salience Theory there exists some $z > 0$ such that $f_{0,z} \prec f_{\alpha,z}$. This representation sharply contrasts with standard economic theory, specifically with the Sure-Thing Principle of Subjective Expected Utility (SEU) Theory (Savage, 1954). For SEU, the addition of z as portrayed necessarily does not change preferences.

All of that considered, the value function that represents the DM's preferences and the effect of Salience in their choice is given by:

$$V(f_i) = \sum_{s \in S} p_s \omega(\mathbf{x}_s) v(x_s^i)$$

where $\omega(\mathbf{x}_s)$ is a salience function applied to state s, given the choice and consequence sets. $v(x_s^i)$ is the value of the outcomes associated with the choice of act i if state s is true. p_s is the probability of state s happening. Since Bordalo et al. (2012) proposes a theory of choice under risk - not yet considering the case of decision under uncertainty or ambiguity - they assume that p_s are objective and known probabilities that exist for each state $s \in S$. In our model, we build upon the relaxation of that assumption.

4 The Model

4.1 Main Assumptions

Next we describe two main assumptions of our model. The first one concerns the representation of preferences over constant acts - i.e., acts that do not involve risk and result in the same outcome no matter what is the true state of the world. The assumptions assert that there is a value function that represents the DM's preferences over such acts, with the standard properties of other similar models.

(A1 - Value Function on Acts) Let c_{x_i}, c_{x_j} be constant acts, with consequences $x_i, x_j \in X$ for any $s \in S$, respectively. Then, there exists a value function $v : \mathbb{R} \to \mathbb{R}$, continuous, strictly increasing, and normalized so that $v(0) = 0$ such that, $c_{x_i} \succeq c_{x_j}$ if and only if $v(c_{x_i}) \geq v(c_{x_j})$.

Our second assumption states that the choice problem analyzed with our problem is not trivial, i.e., there is at least one act that is not constant in the choice set.

(A2 - Nontriviality) Define a non-constant act f as one where $f(s_i) \neq f(s_j)$ for some non-null $s_i, s_j \in S$, and a not-certain state of the world as $s \in S$ such that $\mu(\Pi_i(s)) < 1$ for some non-null $\Pi_i \in \Pi$. Then, we assume there exists some non-constant act $f \in F$, with outcome contingent on at least one not-certain state of the world $s \in S$.

With that taken care of, we can move to the main definition of our model in the next section.

4.2 Ambiguity Adjustment Function

Given our preliminary setting, we define a function Ψ that represents how probability weighting representing preferences may be affected by event ambiguity[5].

Definition 3 (Ambiguity Adjustment Function Ψ): Let S be a finite discrete state-space and $\mathcal{A}(S)$ its Borel σ-algebra. Let $f_n : S \to X$ be acts, where $X \in \mathbb{R}$ is an arbitrary set of consequences, $F = \{f_1, ..., f_N\}$ is a choice set and $\mathcal{P}^n$ is the partition of S induced by some act $f_n \in F$. Let $\Delta(S)$ be the set of all probability measures on $(S, \mathcal{A}(S))$, μ a probability distribution measure on $\Delta(S)$, $\mu : \Delta(S) \to \mathbb{P}$ with finite variance, and $\Pi = \{\pi_i \in \Delta(S) : \mu(\pi_i) > 0\}$ the set of non-null (first-order) probability measures. Then, $\Psi : \mu \to [0, 1]$ is an Ambiguity Adjustment Function if the following properties hold:

P1(Degenerate Distribution Independence) For any non-null event $E \in \mathcal{A}(S)$, $\lambda \in (0, 1)$, degenerate μ_1, μ_2 and distribution measures μ_i, $\Psi(\mu_1(E)) \geq \Psi(\mu_2(E))$ if and only if
$\Psi(\lambda\ \mu_1(E) + (1 - \lambda)\ \mu_i(E)) \geq \Psi(\lambda\ \mu_2(E) + (1 - \lambda)\ \mu_i(E))$.

Moreover, under any distribution μ, for any null event E_0, $\Psi(\mu(\Pi(E_0))) = 0$. And, for the universal event $E_S = S$ $\Psi(\mu(\Pi(E_S))) = 1$.

P2(Comparative Second-Order Probabilities) For any non-null event $E \in \mathcal{A}(S)$, all $\lambda, \lambda' \in (0, 1)$ and degenerate distributions $\mu_1, \mu_2, \mu_3, \mu_4$ such that $\Psi(\mu_1(E)) > \Psi(\mu_2(E))$ and
$\Psi(\mu_3(E)) > \Psi(\mu_4(E))$, then $\Psi(\lambda\ \mu_1(E) + (1 - \lambda)\ \mu_3(E)) \geq \Psi(\lambda'\ \mu_1(E) + (1 - \lambda')\ \mu_3(E))$ if and only if $\Psi(\lambda\ \mu_2(E) + (1 - \lambda)\ \mu_4(E)) \geq \Psi(\lambda'\ \mu_2(E) + (1 - \lambda')\ \mu_4(E))$

P3(Small-Distribution Continuity) For all μ_i, μ_j, μ_k satisfying $\Psi(\mu_i(E)) > \Psi(\mu_j(E))$ for an event $E \subset S$, then there is $\lambda \in (0, 1)$ such that:
$\Psi(\mu_i(E)) > \Psi(\lambda\ \mu_k(E) + (1 - \lambda)\ \mu_j(E))$ and $\Psi(\lambda\ \mu_k(E) + (1 - \lambda)\ \mu_i(E)) > \Psi(\mu_j(E))$.

P4(Sub-partition Invariance) For every act-induced partition $\mathcal{P}^i = \bigcup_{i=1}^{n} E_i$ induced by some act f_i, and some fixed μ, $\sum_i \Psi(\mu(\bigcup_{E \in E_i} E)) = \Psi(\mu(E_i))$.

P5(Belief Symmetry) Let $\mathcal{P}^i = \bigcup_{i=1}^{n} E_i$ and $\mathcal{P}^j = \bigcup_{j=1}^{n} E_j$ be act-induced partitions for some acts f_i and f_j, respectively. Then, the following holds:

(i) For all events $E, E' \in \mathcal{P}^i$, $\mu(E) = \mu(E') \implies \Psi(\mu(E)) = \Psi(\mu(E'))$.

(ii) For every event $E \in \mathcal{P}^i \bigcap \mathcal{P}^j$ and $|\mathcal{P}^i| = |\mathcal{P}^j|$ (where $|A|$ denotes the cardinality of a set A), there is a unique $y \in \mathbb{R} : \Psi(\mu(E)) = y$.

[5]Definition 3 is valid for finite discrete state-spaces S. To an adaptation of this definition for continuous state-spaces, see Appendix C.

Properties P1, P2 and P3 are analogous to the Ordinal Event Independence, Comparative Probability and Small-Event Continuity postulates of Subject Expected Utility Theory (Karni, 2014, p.12). P1 states that Ψ has to preserve the ordering of events weighting for any two degenerate distribution measures μ_1 and μ_2, when they are mixed with a non-null distribution measure. P2 states that Ψ should preserve the comparative ranking of ambiguity aversion between two mean-preserving spreads of second-order probability distributions. P3 introduces a form of continuity of probability weighting Ψ on the second-order distribution μ.

Property P4 makes sure that making further partitions (i.e., refinements of the Act-Induced Partition) of an event associated with an outcome x for some act f do not change the weighting (represented by Ψ) of that outcome within this act, i.e., it doesn't change the valuation of act f. In other words, only event partitions that matter to the outcome x or to the description of the state-space S are relevant to probability weighting. Finally, P5 makes sure that that the Ψ function is consistent in that: (i) within an act, if the information about the compound probability of an event of any two events is the same, their weighting is equal; (ii) if the same event is relevant to determine the results of two distinct acts f_i, f_j and the Act-Induced Partitions of the acts are similar (in that they fave the same cardinality), then their weighting is the same for both acts.

4.2.1 Illustrating Properties of the function

We proceed to give examples that illustrate why each postulate is important, and its' consequences to the representation of choice under ambiguity and preferences in this circumstances of decision-making. Take another Ellsberg-urn like example. Suppose there is an urn with 12 balls. Then, define the following degenerate second-order distributions of the urn: 4 red balls, 1 yellow ball and 7 black balls (4R, 1Y, 7B) as μ_1; 1 red ball, 4 yellow balls and 7 black balls (1R, 4Y, 7B) as μ_2.

Given the payoff structure in the table below, it is reasonable to assume that, in a decision-making situation where μ_1 is true, $f_1 \succeq f_2$ - and that happens because the event E_r has more weight in the decision (and so, the favorable outcome for f_1 when E_r is true prevails). This is represented in our model by a higher probability weighting $\Psi(\mu_1(E_r))$. Analogously, in another decision-making situation where μ_2 is true, it is reasonable to assume that $f_2 \succeq f_1$. Combining the preferences of both decision problems, $\Psi(\mu_1(E_r)) \geq \Psi(\mu_2(E_r))$.

	Red (E_r)	Yellow (E_y)	Black (E_b)
f_1	100	0	0
f_2	0	100	0

Table 4: An Ellsberg-urn example. We call this Example 3.

Now, define an additional ball composition as μ_i: the DM knows there is between 4 and 6 red balls, between 4 and 6 yellow balls, and between 0 and 4 black balls ([4,6]R, [4,6]Y, [0,4]B). Now, consider the problem where the probability distribution is $\mu_{i,1} = ([4,5]R, [3,4]Y, [4,6]B)$. That means that we can represent $\mu_{i,1}$ as a convex combination of previous distributions: $\mu_{i,1} = \frac{1}{2}\mu_1 + \frac{1}{2}\mu_i$. Analogously, define $\mu_{i,2} = ([3,4]R, [4,5]Y, [3,5]B) = \frac{1}{2}\mu_1 + \frac{1}{2}\mu_i$.

Under $\mu_{i,1}$, the probability of E_r being true is at least 4/12 (and it may be 5/12 with positive probability) - while under μ_1 the probability of E_r being true was precisely 4/12. Therefore, it is reasonable to assume that, if under μ_1 $f_1 \succeq f_2$, under $\mu_{i,1}$ also $f_1 \succeq f_2$ - since the event that gives a favorable outcome for f_1 may be now more probable than in the previous situation, favoring the act f_1 even more. Moreover, since the only change that happened was in the compound probability distribution of events, it is reasonable to assume that this means that $\Psi(\lambda\mu_1(E_r) + (1-\lambda)\mu_i(E_r)) \geq \Psi(\mu_1(E_r))$. Analogously, under $\mu_{2,i}$, it is sensible to presume that $f_2 \succeq f_1$ still holds - as it happened under μ_2.

Because of this, it is reasonable to assume that the weighting function in such situations should be such as $\Psi(\mu_{1,i}(E_r)) \geq \Psi(\mu_1(E_r)) \geq \Psi(\mu_{2,i}(E_r)) \geq \Psi(\mu_2(E_r))$, where the two weighting functions to the left represent the decision-making problems where $f_1 \succeq f_2$. Rearranging the inequalities, $\Psi(\mu_1(E_r)) \geq \Psi(\mu_2(E_r))$ and $\Psi(\lambda\ \mu_1(E_r) + (1-\lambda)\ \mu_i(E_r)) \geq \Psi(\lambda\ \mu_2(E_r) + (1-\lambda)\ \mu_i(E_r))$, as the postulate P1 poses.

As for P2, the postulate makes sure that the ordering of the values of the function Ψ is independent on the specific compound probability distributions, when these distributions are degenerate. For concreteness, suppose the same setting of Example 3, but the following compound distributions: $\mu_1 = (10R, 0Y, 2B); \mu_2 = (8R, 2Y, 2B); \mu_3 = (4R, 2Y, 6B); \mu_4 = (2R, 2Y, 8B)$. It is reasonable to assume the DM weights the state E_r at least as much under μ_1 (since there are 4 red balls in the urn, for sure) than under μ_2. That is interpreted in our model as $\Psi(\mu_1(E_r)) \geq \Psi(\mu_2(E_r))$. Analogously, we can also assume that $\Psi(\mu_3(E_r)) \geq \Psi(\mu_4(E_r))$. Say $\lambda = 3/4$ and $\lambda' = 1/2$. Then, using λ, λ' to construct convex combinations of μ_1, μ_2 we get the following marginal probabilities for E_r: $\mu_{1,2,\lambda} = (3/4\ \mu_1(E_r) + 1/4\ \mu_2(E_r)) = (9.5R)$; $\mu_{1,2,\lambda'} = (1/2\ \mu_1(E_r) + 1/2\ \mu_2(E_r)) = (9R)$. Therefore, it is reasonable to assume that $\Psi(\mu_{1,2,\lambda}) \geq \Psi(\mu_{1,2,\lambda'})$.

That says something about the ordering of the weights of distributions μ: convex combinations that give more weight to the higher(lower) weighted distribution, should be weighted higher(lower). In the example, μ_1 had the higher weighted E_r, and λ gives more weight to that distribution than λ' - that is, in our model, higher Ψ. Analogously, we expect that logic to happen consistently for any μ, so that, if we apply the same logic to μ_3 and μ_4 in the example, $\mu_{3,4,\lambda} = (3/4\ \mu_3(E_r) + 1/4\ \mu_4(E_r)) = (3.5R)$; $\mu_{3,4,\lambda'} = (1/2\ \mu_1(E_r) + 1/2\ \mu_2(E_r)) = (3R)$ imply $\Psi(\mu_{3,4,\lambda}(E_r)) \geq \Psi(\mu_{3,4,\lambda'}(E_r))$.

Small-Distribution Continuity (P3) introduces a form of continuity for the Ψ function. It makes sure that there is no distribution μ_i that is infinitely more (or less) weighted than some other distribution μ_j, insofar as we can always find some convex combination of μ_j and some μ_k that has more (less) weight than μ_i, for some appropriately chosen λ. In the Example 3 setting, let $\mu_i = ([10, 11]R, [0, 1]Y, 1B); \mu_j = ([0, 1]R, [0, 1]Y, 11B)$. It is, then, reasonable to assume that $\Psi(\mu_i(E_r)) > \Psi(\mu_j(E_r))$, since it is certain that μ_i implies there are more red balls in the urn than μ_j. However, if we choose an appropriate distribution, say $\mu_k = ([11, 12]R, [0, 1]Y, 0B)$, and an appropriate coefficient, say $\lambda = 3/4$, we can construct a distribution that has more weight than μ_i for event E_r. An important remark is that the nontriviality assumption, specially in that there is some uncertainty about which is the true event. For example, P3 would not apply for $(12R, 0Y, 0B)$, since it implies it is certain that a red ball is going to be drawn, and the choice does not involve risk. This is also why P3 highlights that the property is valid for any non-null event that has involves some uncertainty of happening, i.e., the event is a subset of the state-space S. Since our purpose is to design a theory to deal with decision under risk and uncertainty, however, this should not be an issue.

We turn now to Sub-partition invariance, property P4. Differently than the first three properties, this one is not analogous to any of the classic postulates of SEU about preferences. What this property makes sure is that partitioning an event $E_i \in \mathcal{P}^i$ into non-null "sub-events" $E_j \in E_i$ such that $\bigcup_{j=1}^{N} E_j = E_i$ does not affect the weighting of outcomes. That is, refining further an event in a way that is not relevant to the outcomes of f_i nor the description of state-space S does not affect preferences. For example, let's take Example 3 setting, with $\mu = ([4, 6]R, [4, 6]Y, [0, 4]B)$. Now, suppose we write the numbers zero or one in the red balls of the urn, even though the payoffs associated with E_r are not changed, so that between 2 and 3 red balls have the number zero on it (E_{r0}), and between 2 and 3 red balls have the number one (E_{r1}) on it. We can then represent the distribution that takes into account this difference as two separate states of the world as $\mu = ([2,3]R_0, [2,3]R_1, [4,6]Y, [0,4]B)$. Since payoffs associated with E_r under μ_1 and associated with E_{r0} and E_{r1} under μ_2 remain the same - so that the refinement of the event E_r into E_{r0} and E_{r1} is irrelevant to the evaluation of the act - we can represent this modified urn's payoffs as in the table below:

	Red ($E_{r,0}$)	Red ($E_{r,1}$)	Yellow (E_y)	Black (E_b)
f_1	100	100	0	0
f_2	0	0	100	0

Table 5: An Ellsberg-urn example. We call this Example 3.

What P4 states is that $\Psi(\mu(E_r)) = \Psi(\mu(E_{r0})) + \Psi(\mu(E_{r1}))$, that is, that Decision-makers only care about this kinds of event composition changes as long as they affect the outcomes associated with them through the acts in the choice set.

Finally, we turn to the Belief Symmetry postulate, P5. It states a particular way in that the Ψ function is consistent in weighting events. P5-(i) presume consistency between different events within an act with the same compound probability of obtaining; and P5-(ii) implies a form of consistency for the weighting of the same event between acts. Let's go back to our running example. Suppose $\mu = (3R, [0,6]Y, [0,6]B, 3G)$. Then, P5-(i) states that $\Psi(\mu(E_r)) = \Psi(\mu(E_g))$ and $\Psi(\mu(E_y)) = \Psi(\mu(E_g))$, which only stands to reason: since their marginal compound probability distribution is the same, their weight in the decision-making process should also be the same. Now, take acts f_0 and f_α of Table 4 example, and note that their act induced partitions are: $\mathcal{P}^0 = \{E_r, E_g, E_{(y \vee b)}\}$, $\mathcal{P}^\alpha = \{E_y, E_g, E_{(r \vee b)}\}$. There is one event E_g that is in both act-induced partitions. Besides, in both cases there are three events in the Act-Induced partitions, so that the amount of relevant events in determining the outcomes is the same in both cases. Therefore, P5-(ii) states that $\Psi(\mu(E_g))$ has the same value when applied to evaluate acts f_0 and f_α. Intuitively, P5 statements put together say that events to which the DM has the exact the same marginal probability distribution information should be weighted equally, if there is an equal amount of possible outcomes for the acts considered. Its the information about the probability of the events and its' comparison to other relevant events that matter for weighting, and not the event itself.

4.2.2 General Evaluation Acts

Having defined our Ψ function, and applying it to generalize Salience Theory to Choice Under Ambiguity, we can now define the DM's valuation of an act $f_i \in F$:

$$V(f_i) = \sum_{s \in \mathcal{P}^i} \Psi(\mu(\Pi(s))) \; \omega(\sigma(\mathbf{x}_s)) \; v(x^i_s)$$

To make sure property P4 holds, we also define $\Psi(\mu(\Pi(s))) = 1/|E| \cdot \Psi(\mu(\Pi(E)))$ for any $s \in E \in \mathcal{P}^i$. That is, if the refinement of an event in a $E \in \mathcal{P}^i$ is not relevant in determining the outcomes of some act f_i, then we can just divide the weights of the event equally among each state $s \in E$.

Proposition 1 below states that we can define our decision under ambiguity model as one with choice under risk (i.e., when μ is a degenerate distribution, in our model's interpretation) as a specific case.

Proposition 1(Existence of Ψ that generalizes choice under risk) Let S be a finite state-space, and $\mathcal{A}(S)$ its' σ-algebra. Define $\Delta(S)$ as the set of probability measures on the $(S, \mathcal{A}(S))$ space, and $\mu : \Delta(S) \to \mathbb{P}$ a (second-order) probability distribution measure on $\Delta(S)$. Define $\Pi = \{\Pi_i \in \Delta(S) : \mu(\Pi_i) > 0\}$. Suppose μ is degenerate. Then, any Ψ function such that $\Psi(\mu(\Pi(s))) = \pi_s$ for every $s \in S$, where $\pi : S \to \mathbb{P}$ is a probability measure on S. Moreover, $\pi_s = \mathbb{E}_\mu(\mathbb{E}_\Pi(s))$, i.e., π_s is the expected probability of state s being true.

This makes sure that, in situations of decision under risk, we can define a Ψ function that collapses to a regular weighting by the objective probability of each event E relevant to evaluate each act. In other words, this makes sure that our model is a generalization of Salience Theory of Choice Under Risk (Bordalo et al., 2012). An example of Ψ function that satisfies not only postulates P1 through P5 but also an additional postulate P6 - that introduces likelihood insensitivity in our model - is given in section 4.2.3. A simpler - but not that interesting to our purposes - parametric example of function that satisfies only postulates (P1) through (P5) is also given in Appendix B.

Going back to our running example, without loss of generality, by assumption A1 we can normalize the value function so that $v(0) = 0$. Our assumption of the example was that $\Pi = \{\Pi_1, ..., \Pi_7\}$, and $\mu(\Pi_1) = \mu(\Pi_2) = ... = \mu(\Pi_7) = \frac{1}{7}$. We display again below the payoff and the non-null first-order probability distributions tables.

	Red (E_r)	Yellow (E_y)	Black (E_b)	Green (E_g)
f_0	100	0	0	150
f_α	0	100	0	150

Table 6: An Ellsberg-like ambiguity example (Example 1). We also refer to the example of this table as the running example, throughout the text.

	π_r	π_y	π_b	π_g
Π_1	3/12	0	6/12	3/12
Π_2	3/12	1/12	5/12	3/12
Π_3	3/12	2/12	4/12	3/12
Π_4	3/12	3/12	3/12	3/12
Π_5	3/12	4/12	2/12	3/12
Π_6	3/12	5/12	1/12	3/12
Π_7	3/12	6/12	0	3/12

Table 7: Non-null first-order probability distributions in our running example.

We also note that the salience function properties make sure that E_r and E_y are the most salient states, i.e. $\omega_r = \omega_y = \omega$ and $\omega > \omega_b > \omega_g$. Then, the value functions of each act are:

$$
\begin{aligned}
V(f_0) = \Psi(\mu(E_r))\ \omega\ v(100) + \Psi(\mu(E_{y\ \vee\ b}))\ v(0))(\omega + \omega_b) + \Psi(\mu(E_g))\ \omega_g\ v(50) \\
= \Psi(\mu(E_r))\ \omega\ v(100) + \Psi(\mu(E_g))\ \omega_g\ v(150)
\end{aligned}
$$

$$
\begin{aligned}
V(f_\alpha) = \Psi(\mu(E_{r\ \vee\ b})\ v(0))(\omega + \omega_b) + \Psi(\mu(E_y))\ \omega\ v(100) + \Psi(\mu(E_g))\ \omega_g\ v(50) \\
= \Psi(\mu(E_y))\ \omega\ v(100) + \Psi(\mu(E_g))\ \omega_g\ v(150)
\end{aligned}
$$

By our running example assumptions about μ, and by Belief Symmetry (Property P5), $\Psi(\mu(E_g))$ is equal in both equations. Therefore, in our model the DM's choice hinge on whether $\Psi(\mu(E_r)) \geq \Psi(\mu(E_y))$ - in which case $f_0 \succeq f_\alpha$ - and/or $\Psi(\mu(E_r), |S|) \leq \Psi(\mu(E_y), |S|)$, which implies $f_0 \preceq f_\alpha$.

By Proposition 1, we know that we can define a Ψ function such that $\Psi(\mu(E_r)) = \mathbb{E}_\mu(E_r)$. Since $\mathbb{E}_\mu(E_r) = \mathbb{E}_\mu(E_y)$, to determine the DM's preference we have to determine whether $\Psi(\mu(E_y)) \leq \mathbb{E}_\mu(E_y)$ (i.e., the DM is ambiguity averse) and/or $\Psi(\mu(E_y)) \geq \mathbb{E}_\mu(E_y)$ (i.e., the DM is ambiguity seeking). Based in our model, we give a possible answer to that question in the next section.

4.2.3 Additional Property P6 and Ambiguity Attitudes

We take advantage of the growing evidence on ambiguity attitudes to add a property to Ψ to make sure the model is a good description of the stylized facts about Decision Under Ambiguity (Trautmann & van de Kuilen, 2015). From the experimental and empirical literature, we know that ambiguity attitude is a function of: (i) the "likelihood" of the event; (ii) whether the outcome associated to that event for some act is in the gain or loss domain; (iii) how many possible states of nature define the outcomes of each act.

We represent those stylized facts by postulate P6, stated below:

Definition 3.1 (Contextual Ambiguity Adjustment Function Ψ) Let postulates **P1-P5** of Definition 3 hold for a Ψ function. Then, Ψ is a Contextual Ambiguity Adjustment Function if postulate P6 also holds:

P6(Partition Monotonicity) For any non-degenerate probability distribution μ_1 and μ_2 such that $\mu_1(\Pi(E_i)) = \mu_2(\Pi(E_i))$ for some $E_i \in \mathcal{P}^i$ and $\mu_1(\Pi(E_j)) = \mu_2(\Pi(E_j))$ for some $E_j \in \mathcal{P}^j$, if $|\mathcal{P}^i| < |\mathcal{P}^j|$, then $\Psi(\mu_1(\Pi(E_i))) \geq \Psi(\mu_2(\Pi(E_i)))$ implies $\Psi(\mu_1(\Pi(E_j))) \geq \Psi(\mu_2(\Pi(E_j)))$.

To illustrate what property P6 assures us, consider our running example, but now assume $\mu_1 = ([0,6]R, [0,6]Y, 3B, 3G)$ and $\mu_2 = ([0,6]R, [0,6]Y, 6B, 0G)$. That is, under μ_1, the DM knows there are 3 black balls and 3 green balls, an knows that the remaining balls are either red or yellow, in unknown proportion. Under μ_2, the DM knows there are 6 black balls, and also knows the remaining balls are either red or yellow, in unknown proportion. If the DM interprets the unknown proportion as an equal probability of each combination of red and yellow balls being true, then it is easy to see that the marginal probability distribution of event E_y being true is the same in μ_1 and μ_2. Analogously, the same is true for event E_b. Therefore, $\mu_1(\Pi(E_r)) = \mu_2(\Pi(E_r))$ and $\mu_1(\Pi(E_y)) = \mu_2(\Pi(E_y))$. However, P6 proposes that likelihood insensitivity affects weighting differently under each distribution, in the following way: under μ_2, the DM recognizes only three possible (non-null) events (E_r, E_y and E_b), and therefore the naive probability that he considers when balancing the weighting of each event is 1/3; on the other hand, under μ_1, there are 4 possible non-null events, and thus the naive probability that he considers when balancing the weighting of each event is 1/4.

For concreteness, we provide below a parametric example of Ψ for a discrete state-space S[6], and we heuristically describe that is actually an ambiguity adjustment function[7] and then apply it to the example to draw the same conclusion as in the previous paragraph. For simplicity, we also drop the function argument in parenthesis when dealing with the parametric function below, from now on.

[6] We discuss continuous state-spaces in Appendix C.

[7] The complete proof is in the Appendix.

$$\psi(\mu(\Pi(E))) = \mathbb{E}_\mu[E]^{(1-\gamma\cdot\sigma_\mu(E^i_s))} \left(\frac{1+\eta}{|S^i|}\right)^{\gamma\cdot\sigma_\mu(E^i_s)} \tag{1}$$

where $\sigma_\mu(E^i_s) = \sigma_\mu(\Pi(E^i_s))$ is the standard deviation of the marginal probability distribution of a $\mu(\Pi(E^i_s))$. $\mathbb{E}_\mu(s) = \mathbb{E}_\mu(\Pi(s))$ is the expected probability of state s implied by the distribution $\mu(\Pi(s))$, $S^i = \{s_1, ..., s_n\}$ is the subset of S for which f_i is well-defined, and $|S^i|$ its' cardinality. We assume $\gamma \in [0, 2]$ $\eta \in (-1, 1)$. Finally, following postulate P4, to calculate ψ for states $s \in S$ (rather than events $E^i_s \in S$ that may contain multiple states), we consider $\psi(\mu(\Pi(s))) = 1/|E^i_s|\psi(\mu(\Pi(E)))$, for any $s \in E^i_s \subseteq \mathcal{P}^i$ and act f_i, where $|E^i_s|$ is the cardinality of such event.

Here, the $1/|E^i_s|$ term is used just to make sure that, for states in the same event of an act-induced partition, the weighting is the same. For instance, in Example 1, we want s_y and s_b to be equally weighted (each one with half the weighting of the event $E_{y \vee b}$), but that the information about the probability distribution of $E_{y \vee b}$ to be considered . Then, the first term in the outer parenthesis indicates the part of weighting that is based on the information about probability distributions available to the DM ($\mathbb{E}_\mu[s]$ and $\sigma_\mu(E^i_s)$), while the second term indicates how other information about the choice set F itself (such as $|S^i|$) may distort that weighting, specifically through likelihood insensitivity.

Since the standard deviation of $\Pi(E)$ is always positive and bounded above at $\frac{1}{2}$[8], and we assume $\gamma \in [0, 2]$ the exponents of the equation are always between zero and one. γ can be seen as the degree to which the DM "distorts" the weight given to an event as a function of the imprecision of the information about its' probability.

By our nontriviality assumption $A2$, for any ambiguous event $|\mathcal{P}^i_s| \geq 2$, which means that $|S^i| \geq 2$, i.e., there are at least two possible events that affect the outcome of f_i and form a partition of S. Moreover, $|S^i|$ is constant for each state $s \in S^i$ [9]. That means $\eta \in (-1, 1)$ assures that the second term in parenthesis in 4.2.3 is always greater than zero, but less than one. Here, η is the parameter that indicates at what value of expected likelihood $\hat{\mathbb{E}}_\mu[s]$ the DM is indifferent between being totally ignorant about the probability distribution (and relying heavily on a naive distribution to determine the weight of state s in the evaluation of the act) and knowing for sure that the probability of state s happening is $\pi(s) = \hat{\mathbb{E}}_\mu[s]$. Higher η indicate a higher indifference point $\hat{\mathbb{E}}_\mu[s]$. Also, for values close to the extremes in the $(-1, 1)$ range, this indifference point may not exist at all under some distributions μ.

Proposition 2: The ψ function represented in equation is a Contextual Ambiguity Adjustment Function, that is, it satifies postulates P1 through P6.

Proof: in Appendix A.

To make it tangible, take a modified version of our running example. Let there be an Ellsberg urn with 100 balls, and consider two possible information sets: (1) the DM knows that the Ellsberg urn contains either zero or $\bar{r}$ red balls, and the remaining balls can be either, yellow, black or green, in unknown proportion; (2) the DM only knows that the urn contains either $\underline{r}$ or 100 red balls, and

[8] To see this, note that the maximum standard deviation for $\Pi(E)$ is obtained when the probability mass is concentrated in its' extreme points, since $\Pi(E) \in [0, 1]$ is bounded. That is, the maximum standard deviation of $\Pi(s)$ is obtained when $\pi_1 = 0, \pi_2 = 1$ and $\mu(\pi_1) = 1/2, \mu(\pi_2) = 1/2$. In that case, $\max(\sigma_\mu(\Pi(s))) = \sqrt{1/2(1 - \mathbb{E}_\mu(\Pi(s)))^2 + 1/2(0 - \mathbb{E}_\mu(\Pi(s)))^2} = 1/2$

[9] There may be some instances where, for some event $E \in \mathcal{P}^i_s \cap \mathcal{P}^j_s$, $|\mathcal{P}^i_s|$, when the partition of $(\mathcal{P}^i_s \setminus E)$ differs from $(\mathcal{P}^j_s \setminus E)$, which would violate the Belief Symmetry property (P5). Hovever, economically significant examples such as these are rather uncommon, both in the experimental and empirical literature.

the remaining balls can be either, yellow, black or green, in unknown proportion[10]. In our model, assuming the DM interprets that each possible red ball composition as being equally likely, Case 1 would be described by $\Pi_1(E_r) = 0, \Pi_2(E_r) = \bar{r}/100$ and $\mu_1(\Pi_1) = \mu_1(\Pi_2) = 1/2$. Similarly, Case 2 would be described by $\Pi_3(E_r) = \underline{r}, \Pi_4(E_r) = 1$ and $\mu_2(\Pi_3) = \mu_2(\Pi_4) = 1/2$.

Also, note that in case (1), the expected probability of a drawn ball being red is $\bar{r}/2 \in [0, 0.5]$, while in case (2) this expected probability is $(0.5 + (1 - \underline{r})/2) \in [0.5, 1]$. Figure 1 shows the effect of varying η and γ in the described cases.

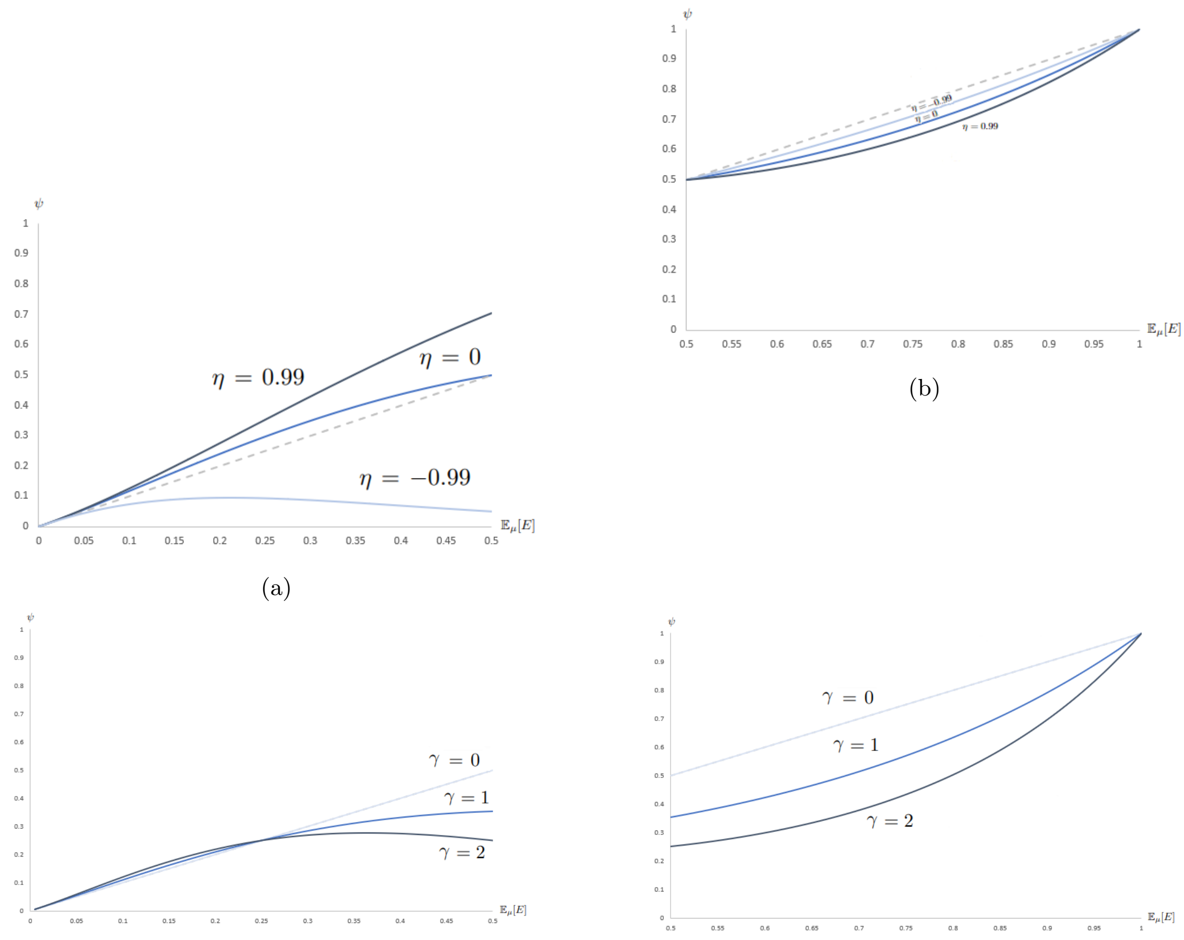
The Figure shows that greater values of γ indicate greater distortion in probability weighting due to ambiguity. In other words, it means that the DM is more sensitive to ambiguity for greater γ, i.e., the difference in weighting is larger for the DM, given an ambiguity level represented by the standard deviation of the second-order distribution $\mu(\Pi(E))$. γ alone does not predict nor indicate ambiguity attitude, except when $\gamma = 0$, in which case the DM is ambiguity neutral for any ambiguous event.

As we previously mentioned, η indicates where the ψ function crosses the 45°curve, so that higher η dislocate that point to the right in the graph. Also, it may be that, for some extreme values of η and an act f_i that implies S^i, the two curves cross only at $\mathbb{E}_\mu[E] = 0$ and $\mathbb{E}_\mu[E] = 1$. In subfigures (a) and (b) of Figure 1, we can see that for $\eta = -0.99$.

We now go back to our running example, and see how act evaluation would be considered under the parametric ψ function stated. In that case, the value functions of the lotteries would be:

$$
\begin{aligned}
V(f_0) = \mathbb{E}_\mu[E_r]^{(1-\gamma\cdot\sigma_\mu(E_r))} \left(\frac{1+\eta}{|S^0|}\right)^{\gamma\cdot\sigma_\mu(E_r)} \quad \omega\ v(100) \\
+\left(\frac{1}{|E^i_{y\ \vee\ b}|} + \frac{1}{|E^i_{y\ \vee\ b}|}\right) \mathbb{E}_\mu[E_{y\ \vee\ b}]^{(1-\gamma\cdot\sigma_\mu(E_{y\ \vee\ b}))} \left(\frac{1+\eta}{|S^0|}\right)^{\gamma\cdot\sigma_\mu(E_{y\ \vee\ b})} \quad (\omega + \omega_b)\ v(0) \\
+\mathbb{E}_\mu[E_g]^{(1-\gamma\cdot\sigma_\mu(E_g))} \left(\frac{1+\eta}{|S^0|}\right)^{\gamma\cdot\sigma_\mu(E_g)} \\
= (\frac{3}{12})^{(1-\gamma\cdot 0)} \cdot \left(\frac{1+\eta}{4}\right)^{\gamma\cdot 0} \quad \omega\ v(100) + \left(\frac{1}{2} + \frac{1}{2}\right) (\frac{6}{12})^{(1-\gamma\cdot 0)} \left(\frac{1+\eta}{4}\right)^{\gamma\cdot 0} \quad (\omega + \omega_b)\ v(0) \\
+(\frac{3}{12})^{(1-\gamma\cdot 0)} \cdot (\frac{1+\eta}{4})^{\gamma\cdot 0}\ \omega_g\ v(150) \\
= \frac{1}{4}\ \omega\ v(100) + \frac{1}{4}\ \omega_g\ v(150)
\end{aligned}
$$

[10]In both cases, we assume that each possible ball color (red, yellow, black or green) is relevant relevant in determing the outcome of the evaluated act.



(a) (b) (c) (d)

Figure 1: Effect of varying parameters η and γ on ψ, as a function of the expected probability of the ambiguous event. Subfigures (a) and (b) show the variation of η (with fixed $\gamma = 1$) in Cases 1 and 2, respectively. Subfigures (c) and (d) show the variation of γ (with fixed $\eta = 0$) in Cases 1 and 2, respectively. Cases 1 and 2 are described in the text, in this section. Finally, the gray dashed line represents the 45°line, that is the weighting f ψ for unambiguous events.

$$
\begin{aligned}
V(f_\alpha) = \mathbb{E}_\mu[E_y]^{(1-\gamma\cdot\sigma_\mu(E_y))} \left(\frac{1+\eta}{|S^\alpha|}\right)^{\sigma_\mu(E_y)} \quad \omega\ v(100) \\
+\left(\frac{1}{|E^i_{r\ \vee\ b}|}+\frac{1}{|E^i_{r\ \vee\ b}|}\right)\mathbb{E}_\mu[E_{r\ \vee\ b}]^{(1-\gamma\cdot\sigma_\mu(E_{r\ \vee\ b}))}\left(\frac{1+\eta}{|S^\alpha|}\right)^{\gamma\cdot\sigma_\mu(E_{y\ \vee\ b})} \quad (\omega+\omega_b)\ v(0) \\
+\mathbb{E}_\mu[E_g]^{(1-\gamma\cdot\sigma_\mu(E_g))}\left(\frac{1+\eta}{|S^\alpha|}\right)^{\gamma\cdot\sigma_\mu(E_g)} \\
=\left(\frac{3}{12}\right)^{(1-\gamma/3)}\cdot\left(\frac{1+\eta}{4}\right)^{\gamma/3}\ \omega\ v(100)+\left(\frac{1}{2}+\frac{1}{2}\right)\left(\frac{6}{12}\right)^{(1-\gamma/6)}\left(\frac{1+\eta}{4}\right)^{\gamma/6}\ (\omega+\omega_b)\ v(0) \\
+\left(\frac{3}{12}\right)^{(1-0)}\cdot\left(\frac{1+\eta}{4}\right)^{0}\ \omega_g\ v(150) \\
=\left(\frac{1}{4}\right)^{(1-\gamma/3)}\cdot\left(\frac{1+\eta}{4}\right)^{\gamma/3}\ \omega\ v(100)+\frac{1}{4}\ \omega_g\ v(150)
\end{aligned}
$$

Considering $\eta \in (-1,1)$ and $\gamma \in [0,2]$, then for any $\gamma > 0$, $V(f_\alpha) < V(f_0)$ and the DM chooses act f_0 over f_α, which is the same ambiguity aversion result before in the nonparametric application. If $\gamma = 0$, i.e., the individual is ambiguity neutral for any event, then $V(f_\alpha) = V(f_0)$ and the DM is indifferent between f_0 and f_α.

5 Ambiguity Attitudes

One of the main reasons that make decision under ambiguity an interesting topic is the fact that it is usual that people differentiate between acts that are contingent on unambiguous and ambiguous events - and also between distinct degrees of ambiguity of the events. We can say heuristically that the way people relate the imprecision of information on likelihoods of events to their choices is what we can call ambiguity attitude. So, when comparing acts with the same payoff structure, we can derive ambiguity attitude if payoffs are contingent on events that have the same expected probability of happening, but with different levels of imprecision (or vagueness) of the information about that probability with different levels of precision. In our running example 1 that is exactly what we are analyzing, that has the payoff \$100 contingent on E_r for f_0, and E_y for f_α. Therefore, for the \$100 payoff and the $F = \{f_0, f_\alpha\}$ choice set, we can say that the individual is ambiguity averse if $f_0 \succeq f_\alpha$, and ambiguity seeking if $f_0 \preceq f_\alpha$. If both are true (i.e., $f_0 \sim f_\alpha$), then the DM is ambiguity neutral. In different settings, both ambiguity aversion and ambiguity seeking behavior have been observed, which is why we use the more appropriate term ambiguity attitude (Trautmann & van de Kuilen, 2015).

Thus, we argue that ambiguity attitude is a result of two factors: (i) the (im)precision of available information about the probability distribution of events that are relevant to the outcomes of any act in the choice set; (ii) how people interpret this and incorporate other information about the choice set to determine their preferences.

In our model, (i) is given by $\mu(\Pi)$, that attaches (second-order) probabilities to each plausible probability distribution of events. (ii) is described by the Ψ function that relates how that information - and complementary information about the choice set - is incorporated in the DM's choice. We discuss this in more detail in the next subsections.

Our model puts forward an axiomatic approach about how people interpret probabilities within the context of choice under ambiguity. With that approach, we are able to make predictions that match stylized facts of the literature and relates them in a meaningful way. That contrasts with previous models, that require an additional assumption about what ambiguity attitude the DM has based on the model parameters, such as the Smooth Ambiguity Preference Model (Klibanoff et al., 2005) and the α-maxmin model (Gilboa & Schmeidler, 1989). Our model closely relates to the idea of the Gajdos et al. (2008) model that indicates that people are in general averse to information imprecision. However, we incorporate all the information about first and second order probability distributions into the DM's choice, and not only the most pessimistic scenario, in terms of utility. That is, we specifically indicate how the DM's ambiguity attitude varies as a function of the expected probability of a given event and characteristics of the presented choice set. Other second-order belief models in the literature, in contrast, consider only outcomes (and the value or utility function that represents the DM's evaluation of outcomes) and an assumed criterion that assumes the same transformation of information about probabilities into weights in evaluating an act, independent of the high or low likelihood of an event or other characteristics of the choice set. We expand on our model's interpretation on how this additional factors influence ambiguity attitude in the next subsections.

5.1 Ambiguity Attitudes and the Choice Set

As it happens with Salience Theory of Choice Under Risk, our model also incorporates the context of a decision, interpreted as the influence of choice set characteristics in the DM's choice. STR assimilates how the state-contingent outcomes and the contrast of that outcomes across acts may affect a DM's choice, as exemplified in our lottery example in Section 3.2. We now incorporate the amount of possible states of the world to the context of Decision Under Ambiguity, represented by the cardinality of the state-space ($|S^i|$) in our model, for discrete state-spaces. This addition makes our model able to explain the widely reproduced likelihood insensitivity effect: people do not sufficiently discriminate between different levels of likelihood of an ambiguous event, transforming subjective likelihood towards a naive equal distribution among events (Dimmock et al., 2012).

Therefore, it is useful to see how changing the cardinality of the choice set may change the DM's choice in a modified version of our running example. Suppose now we get a second urn, that is equal to the one in our running example, except that we substitute the green from our running example for black balls, and inform the DM about that. All the remaining assumptions remain the same as in our running example. From now on, we call this second urn the $(-g)$ urn.

Consequently, now the DM knows for sure that there are no green balls in the $-g$ urn, and now any payoffs contingent on the ball drawn from this second urn being green ($E_{g,-g}$) become meaningless. We represent the new state-space, that considers both urns as $S = \{E_r, E_y, E_b, E_g, E_{r,-g}, E_{y,-g}, E_{b,-g}\}$, where all events with the $-g$ indicate events associated with the ball drawn from the second urn. We also make the following changes in the choice set: (i) add a new act to the choice set ($f_{\alpha,-g}$), contingent on the ball drawn from the second urn; (ii) modify the acts in our running example, so that both pay \$0 when a green ball is drawn from the original urn. We call this modified version of the acts $f_{0,x_g=0}$ and $f_{\alpha,x_g=0}$[11]. Thus, the choice set is now $F = f_{0,x_g=0}, f_{\alpha,x_g=0}, f_{\alpha,-g}$. The payoff matrix of the acts is then given below:

With this new set of information about the second urn, the DM also knows that there is between 3 and 9 black balls in the urn. The plausible probabilities of $E_{r,-g}$ and $E_{y,-g}$ remain the same and $E_{g,-g}$ is a null-event. So, the set Π_{-g} of non-null first-order distributions for the second urn are:

It is easy to see that $f_{\alpha,-g}$ outcomes are contingent on ambiguous events, since she has somewhere

[11] This example can be seen as an adaptation of a thought experiment due to Takashi Hayashi (Ahn, 2008), that - to the best of our knowledge - may be one of the first ones to explicitly pose the problem of how probability estimates of relevant events depend on the choice set.

	Red E_r	Yellow E_y	Black E_b	Green E_g	Red - Urn $-g$ $E_{r,-g}$	Yellow - Urn $-g$ $E_{y,-g}$	Black - Urn $-g)$ $E_{b,-g}$
$f_{0,x_g=0}$	100	0	0	0			
$f_{\alpha,x_g=0}$	0	100	0	0			
$f_{\alpha,-g}$					0	100	0

Table 8: Example 1 - an Ellsberg-like ambiguity example.

	π_r	π_y	π_b
$\Pi_{1,-g}$	3/12	0	9/12
$\Pi_{2,-g}$	3/12	1/12	8/12
$\Pi_{3,-g}$	3/12	2/12	7/12
$\Pi_{4,-g}$	3/12	3/12	6/12
$\Pi_{5,-g}$	3/12	4/12	5/12
$\Pi_{6,-g}$	3/12	5/12	4/12
$\Pi_{7,-g}$	3/12	6/12	3/12

Table 9: Non-null first-order probability distributions in our running example (Example 4.1).

between 0 and 1/2 probability of winning \$100 and somewhere between 1/2 and 1 probability of getting nothing. Observe that $\mathbb{E}_\mu(\mathbb{E}_\Pi(\pi_y) = \mathbb{E}_\mu(\mathbb{E}_{\Pi_{-g}}(\pi_{y,-g}) = 1/4$, and the marginal distribution $\Pi_{-g}(\pi_y)\ U_d(\Pi)$, both the same as in our main running example. This new example is then constructed exactly in a way that the choice set changes without the first and second-order probability of non-zero outcomes changing, so that we can analyze concretely how the model responds to a change in the cardinality of the relevant partition of the state-space for each act. Note that, since the events in the original urn are uncorrelated with the ones in urn $-g$, now $S^{0,x_g=0} = S^{\alpha,x_g=0} = \{E_r, E_y, E_b, E_g\}$ and $S^{0,-g} = \{E_{r,-g}, E_{y,-g}, E_{b,-g}\}$. For clarity, we represent those partitions in the figure below, for the case that $\Pi = \Pi_4$ (i.e., 1/4 probability for E_r, E_y, E_b and E_g) in the original urn and $\Pi_{-g} = \Pi_{4,-g}$ in the $-g$ urn (i.e., 1/4 probability for $E_{r,-g}$ and $E_{y,-g}$ and probability 1/2 for $E_{b,-g}$).

Applying the DM evaluation function to this new setting, we get the following results:

$$
\begin{aligned}
V(f_{0,x_g=0}) = \Psi(\mu_{E_r}, |S^{0,x_g=0}|)\ \omega\ v(100) + (\Psi(\mu(E_y), |S^{0,x_g=0}|)\ \omega\ v(0) \\
+(\Psi(\mu(E_{b\ \vee\ g}), |S^{0,x_g=0}|)\ \underline{\omega}\ v(0) \\
= \Psi(\mu_{E_r}, |S_{0,x_g=0}|)\ \omega\ v(100)
\end{aligned}
$$

$$
\begin{aligned}
V(f_{\alpha,x_g=0}) = \Psi(\mu_{E_y}, |S^{\alpha,x_g=0}|)\ \omega\ v(100) + (\Psi(\mu(E_r), |S^{\alpha,x_g=0}|)\ \omega\ v(0) \\
+(\Psi(\mu(E_{b\ \vee\ g}), |S^{\alpha,x_g=0}|)\ \underline{\omega}\ v(0) \\
= \Psi(\mu_{E_y}, |S_{\alpha,x_g=0}|)\ \omega\ v(100)
\end{aligned}
$$

$$
\begin{aligned}
V(f_{\alpha,-g}) = \Psi(\mu_{E_{y,-g}}, |S^{\alpha,-g}|)\ \omega\ v(100) + \Psi(\mu(E_{(r,-g)\ \vee\ (b,-g)}), |S^{\alpha,x_g=0}|\ \underline{\omega}\ v(0) \\
= \Psi(\mu_{E_{y,-g}}, |S^{\alpha,-g}|)\ \omega\ v(100)
\end{aligned}
$$

First, note the salience is the same ω in each state of the final equation (since it envolves a \$100 payoff in one of the acts and null payoffs for the other ones). Then, the DM's choice hinge on the relation

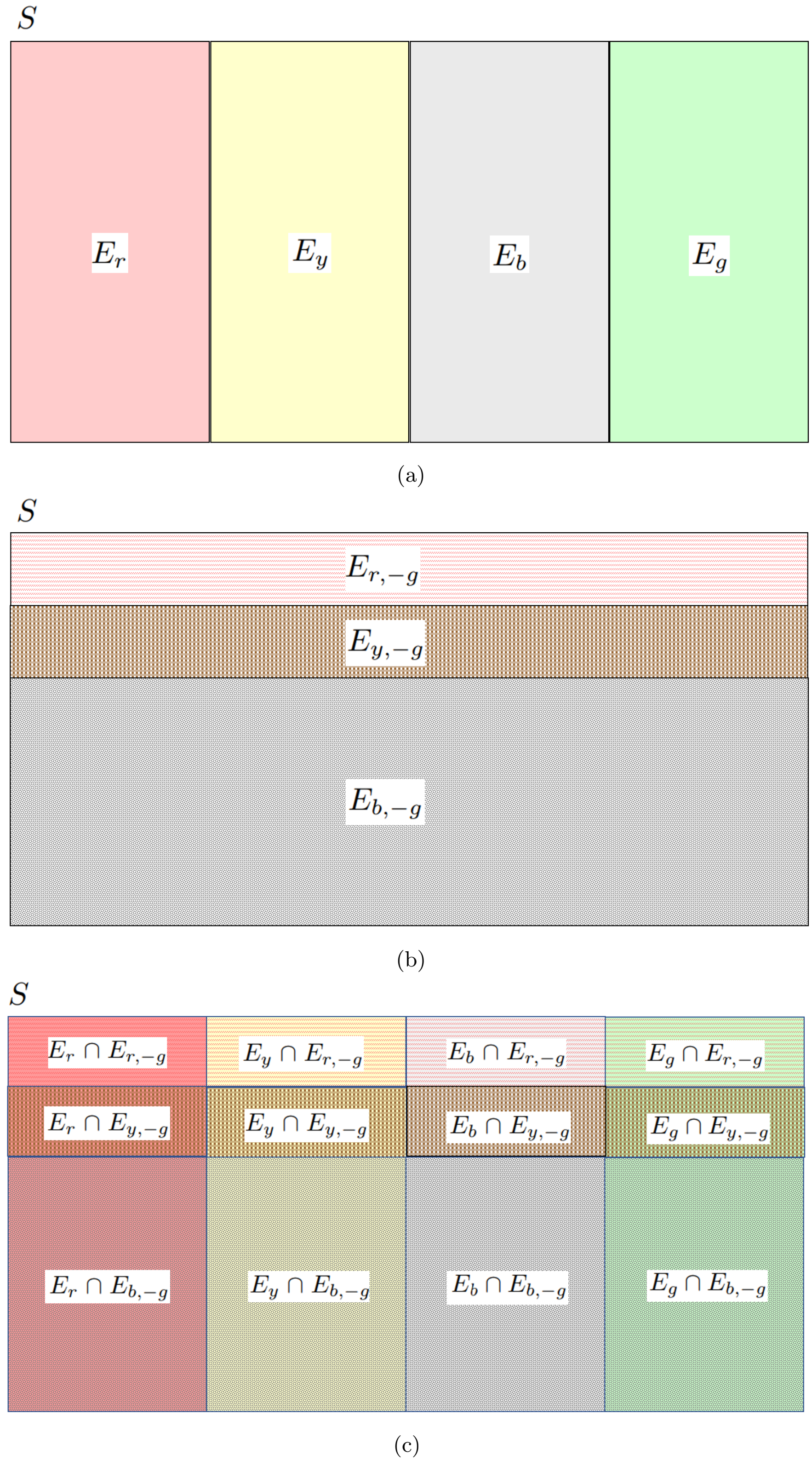


Figure 2: Variations on the partition of the state space in Section 5.1 example, a two-urn Ellsberg decision problem between acts $f_{0,x_g=0}, f_{\alpha,x_g=0}, f_{\alpha,-g}$. Subfigure (a) shows the partition of the state-space S into states that are relevant in determining the outcomes of $f_{0,x_g=0}, f_{\alpha,x_g=0}$, while subfigure (b) shows the relevant partition for the outcomes of $f_{\alpha,-g}$. Since the states in (a) and (b) are uncorrelated, we can represent the entire state-space of the choice problem with the intersections of the events in subfigure (c).

between $\Psi(\mu_{E_r}, |S^{0,x_g=0}|)$, $\Psi(\mu_{E_y}, |S^{\alpha,x_g=0}|)$, $\Psi(\mu_{E_{y,-g}}, |S^{\alpha,-g}|)$. Note that the marginal distributions are such that $\mu_{E_y} = \mu_{E_{y,-g}}$. Then, our property P6 comes into play. Since $|S^{\alpha,x_g=0}| < |S^{\alpha,-g}|$ to determine that $\Psi(\mu_{E_y}, |S^{\alpha,x_g=0}|) \leq \Psi(\mu_{E_{y,-g}}, |S^{\alpha,-g}|)$ and, thus, $f_{\alpha,-g} \succsim f_{\alpha,x_g=0}$. On the other hand, the preference between $f_{\alpha,-g}$ and $f_{0,x_g=0}$ depends on the specific form of the Ψ function, since two effects working in opposite directions are in play: (i) on one hand, $|S^{0,x_g=0}| < |S^{\alpha,-g}|$, which means that the naive distribution that affects weighting for $|S^{\alpha,-g}|$ is larger, by P6; (ii) on the other hand, the marginal distribution $\mu(E_r, |S^{\alpha,-g}|)$ is a mean-preserving spread of $\mu(E_{y,-g}, |S^{\alpha,-g}|)$. By P2, that means that the DM has more precise information about the probability of E_r than $E_{y,-g}$, and then should weight E_r more. Therefore, depending on the sensibility of the DM's weighting to information imprecision and to likelihood insentitivity, the DM's preferences may either be $f_{0,x_g=0} \succsim f_{\alpha,-g}$ or $f_{0,x_g=0} \precsim f_{\alpha,-g}$.

We can observe these relations applying our parametric ψ function to this example:

$$\begin{aligned}\psi(\mu_{E_r}, |S^{0,x_g=0}|) &= \mathbb{E}_\mu[E_r]^{(1-\gamma\cdot\sigma_\mu(E_r))}\left(\frac{1+\eta}{|S^{0,x_g=0}|}\right)^{\gamma\cdot\sigma_\mu(E_r)} \\ &= \left(\frac{1}{4}\right)^{(1-\gamma\cdot 0)}\left(\frac{1+\eta}{4}\right)^{(\gamma\cdot 0)} = \frac{1}{4}\end{aligned}$$

$$\begin{aligned}\psi(\mu_{E_y}, |S^{\alpha,x_g=0}|) &= \mathbb{E}_\mu[E_y]^{(1-\gamma\cdot\sigma_\mu(E_y))}\left(\frac{1+\eta}{|S^{0,x_g=0}|}\right)^{\gamma\cdot\sigma_\mu(E_y)} \\ &= \left(\frac{1}{4}\right)^{(1-\gamma/3)}\left(\frac{1+\eta}{4}\right)^{(\gamma/3)}\end{aligned}$$

$$\begin{aligned}\psi(\mu_{E_{y,-g}}, |S^{\alpha,-g}|) &= \mathbb{E}_\mu[E_{y,-g}]^{(1-\gamma\cdot\sigma_\mu(E_{y,-g}))}\left(\frac{1+\eta}{|S^{0,x_g=0}|}\right)^{\gamma\cdot\sigma_\mu(E_{y,-g})} \\ &= \left(\frac{1}{4}\right)^{(1-\gamma/3)}\left(\frac{1+\eta}{3}\right)^{(\gamma/3)}\end{aligned}$$

First, note that in our parametric function, for any $\gamma \in (0, 2)$, $f_{\alpha,-g} \succ f_{\alpha,x_g=0}$ and $f_{0,x_g=0} \succ f_{\alpha,x_g=0}$ unequivocally. The preference relation between $f_{\alpha,-g}$ and $f_{0,x_g=0}$ than depends on the parameters for aversion to information precision γ and likelihood insensitivity η. Specifically, for any $\gamma > 0$, $\eta \in (-0.25, 1) \implies f_{\alpha,-g} \succ f_{0,x_g=0}$ and $\eta \in (-1, -0.25) \implies f_{0,x_g=0} \succ f_{\alpha,-g}$.

Comparing this with the use of our parametric model in our running example, we can see that now the DM is more prone to overweight $E_{y,-g}$ in the for given η and $\gamma > 0$ parameters, since the last term of $V(f_\alpha)$ is now $(1+\eta)/3)^{\gamma/3}$ instead of $(1+\eta)/4)^{\gamma/3}$. We can interpret that as the DM changing what is his naive probability distribution, now that only three events relevant to the acts are possible - so that the naive probability distribution would be each event occurring with 1/3 chance. That means that a given expected probability of an event is more likely to be perceived as a low-probability event, which has its' weight increased by likelihood insensitivity, in our model.

To sum up, our model considers that the choice set affects probability weighting by the presentation of a set of possible (non-null) states of the world related to the outcome of the bet. Property P6 makes sure that, if there are more possible events associated with the results of a bet, then likelihood insensitivity skews the weighting of an event down (as we saw $\Psi(\mu_{E_y}, |S^{\alpha,x_g=0}|) \leq \Psi(\mu_{E_{y,-g}}, |S^{\alpha,-g}|)$ in our example). That is, the more possible states of the world are presented as relevant to a bet, more this information is interpreted as a lesser probability of any one of the states of the world being true, which is the essence of the likelihood insensitivity effect observed in the literature.

5.2 Ambiguity Attitudes and Second-Order Probabilities

In our model, as it is usual in second-order belief models of decision under ambiguity (Etner et al., 2012), second-order probabilities μ represent the (im)precision of event probability information available to the DM. Since μ is defined on a set of first-order probability distributions of events, the marginal compound probability distribution has finite expected value. Its' standard deviation is taken a measure of the imprecision of the available information. Other than that, μ here is purposely is defined in a broad sense, since many factors may influence the available information about the likelihoods of events, such as related historical data, the framing of the decision problem, and so forth. Furthermore, the model that we present is static, in the sense that it represents a one-shot decision under ambiguity. Eliciting second-order belief formation processes and incorporating dynamic updating of beliefs in our model is an important venue for future research.

We are also able to separate in our model how ambiguity attitudes are affected by events with different expected probability ("likelihood") of happening. Here, we represent the likelihood of an event assessed by the expected probability of the event, given the second-order probabilities μ, for any non-null event. With that, we are able to separate any over/under-weighting of states in decision-making due to lack of information about states and their probability (ambiguity itself, represented by the Ψ function) and due to salience of the known information of outcomes, represented by the salience function ω_s^i and its covariance with the known state-contingent outcomes (Bordalo et al., 2012) - as the STR model already does.

However, we note that second-order distribution by itself does not elicit a DM's ambiguity attitude. Only together with $|S^i|$ and the specific parametric form of Ψ can μ indicate if a state of the world is considered as a highly or lowly likely. It is this comparison of the likelihood of the event with characteristics of the state-space that may pinpoint if a given expected probability of an event is considered high or low - and that in turn imply if likelihood insensitivity causes over or under-weighting of a given state of the world, as we will see in the next section. However, once the low/high likelihood of the event is determined, the properties of our model make sure that low likelihood events are overweighted and high likelihood events are underweighted. Again, this is in agreement with the experimental evidence on likelihood insensitivity (Trautmann & Van de Kuilen, 2015).

It is useful to see how changing the assumptions on μ may change the DM's choice in a modified version of our running example. But now, the DM has the additional information that there are either 0 or 6 yellow balls in the urn, i.e., $\mu^*(\Pi_2) = \mu^*(\Pi_3) = \mu^*(\Pi_4) = \mu^*(\Pi_5) = \mu^*(\Pi_6) = 0$. Now, only Π_1 and Π_7 are plausible (non-null) elements of $\Delta(S)$, i.e., $\Pi = \{\Pi_1, \Pi_7\}$, where $\Pi_1 = (\pi_{E_r} = \frac{3}{12}, \pi_y = 0, \pi_b = \frac{6}{12}, \pi_g = \frac{3}{12})$ and $\Pi_7 = (\pi_{E_r} = \frac{3}{12}, \pi_y = \frac{6}{12}, \pi_b = 0, \pi_g = \frac{3}{12})$. We again assume that the DM assumes that Π_1 and Π_7 are equally likely, that is, $\mu^*(\Pi_1) = \mu^*(\Pi_7) = 1/2$.

Comparing this new μ^* with our running example distribution μ, observe that $\mathbb{E}^*_\mu(\Pi(E_y)) = \mathbb{E}_\mu(\Pi(E_y))$, that is, the expected probability of event E_y is the same in both cases[12], but $\sigma(\mu^*(\Pi(E_y))) > \sigma(\mu(\Pi(E_y))$. However, note that μ and μ^* are non-degenerate distributions, so that property P2 does not apply, even though $\mu^*(\Pi(E_y))$ is a mean-preserving spread of μ_{E_y}. What we can affirm based in our model is that $||\Psi(\mu^*(\Pi(E_y))) - \mathbb{E}_\mu(\Pi(E_y))|| > ||\Psi(\mu(\Pi(E_y))) - \mathbb{E}_\mu(\Pi(E_y))||$, where $||x||$ denote the absolute value of x. That is, $\Psi(\mu^*(\Pi(E_y)))$ is farther from the expected probability of E_y or, in other

[12]In fact, the expected probability is the same for all events considered, but we focus on E_y because it is the relevant one to our conclusions.

words, the probability weighting distortion is larger, even though it is not possible to affirm if it goes in the direction of overweighting or underweighting without further assumptions. So, even though the DM has narrowed the possible first-order probability distributions down to just two alternatives (versus 7 possibilities in our running example), it is not obvious that state E_y becomes more over-weighted (or underweighted) under μ^*. That happens because our model predicts that, given a choice set and two probability distributions with the same expected probability, more information about the second-order probability distribution only increases the weighting of the event if: (i) the expected probability of the event is a certain likelihood insensitivity indifference point (where $\Psi(\mu(E)) = \mathbb{E}_\mu(E)$ for some event E); (ii) the new information imply a decrease in the standard deviation of the second-order probability of the event happening.

We can also observe how that works in our parametric function ψ applied to the example:

$$\psi(\mu(\Pi(E_r))) = \frac{1}{4}$$

$$\begin{aligned}\psi(\mu(\Pi(E_y))) &= \mathbb{E}_\mu[E_y]^{(1-\gamma\cdot\sigma_\mu(E_y))}\left(\frac{1+\eta}{|S^\alpha|}\right)^{\gamma\cdot\sigma_\mu(E_y)} \\ &= \left(\frac{1}{4}\right)^{(1-\gamma/6)}\left(\frac{1+\eta}{4}\right)^{(\gamma/6)}\end{aligned}$$

What this example highlights is that - for a given expected probability of an event - new information only mitigates probability weighting "distortions" (ψ farther from the expected probability of the event) depend on the relationship between sensitivity to information imprecision γ and likelihood insensitivity parameter η. Specifically, for any $\gamma > 0$, $\eta \in (-1, 0) \implies \psi(\mu^*(\Pi(E_y))) < \psi(\mu(\Pi(E_y)))$ and $\eta \in (0, 1) \implies \psi(\mu^*(\Pi(E_y))) > \psi(\mu(\Pi(E_y)))$. There has been recent evidence in favor of underweighting for additional information that increases standard deviation for bets involving gains (Chew et al., 2017), that can be accommodated in our parametric example with $\eta \in (-1, 0)$. However, further empirical and experimental evidence is needed for this to be a stylized fact, and so we construct our model in a way that can predict both behaviours.

5.3 Ambiguity Attitudes and the Outcome Domain

According to the evidence in the literature (Trautmann & van de Kuilen, 2017; other FONTS), one of the factors that influences ambiguity attitude is the outcome domain - interpreted as whether a given bet involves gains or losses, with respect to a reference point. Tipically, this is tested experimentally comparing individuals choices when presented with a choice set involving a risky lotteries of the form xE_00 and an ambiguous lottery involving $xE_\alpha 0$, where E_0 is a risky event (i.e., the objective probability of E_0 happening is known to the decision-maker), and E_α is an ambiguous event. The tests involve tipically the cases where $x > 0$ and $x < 0$ and the same value of expected probabilities for E_0 and E_α. If an individual chooses $xE_\alpha 0$ over xE_00, we conclude that the individual is ambiguity seeking, and, conversely, if she chooses xE_00 over $xE_\alpha 0$ she is ambiguity averse.

A stylized fact drawn from this literature is that there are some regularities in the most common ambiguity attitude behavior of individuals, that are a function of the expected probability of the event for which the outcome of an act is contingent and the outcome domain of the act's results. The ambiguity attitude in each case is portrayed in the table below:

	Domain	
Expected Probability	Loss	Gain
Low	Ambiguity Aversion	Ambiguity Seeking
Mid/High	Ambiguity Seeking	Ambiguity Aversion

Table 10: Ambiguity Attitudes and Effects on Value Function - Unsureness Aversion Theory

One of the main advantages of our model's interpretation of the probability weighting function and the way DM's transform the available information to form their decision weights on events is that we can account for different ambiguity attitudes for different expected probabilities. That is possible because we separate how people interpret information on probabilities of events from other regularities involving how people evaluate acts and outcomes. That has similarities to how Prospect Theory axiomatizes its' probability weighting function (Kahnemann e Tversky, 1979; Prelec, 1998; Wakker, 2010), but here we can make specific predictions about how these behaviors relate to events with different levels of ambiguity, and choice sets that have different implications in terms of how bottom-up attention may affect the DM's choice.

Concretely, by assumption A1 we have that the value function $v(x)$ can be normalized as $v(0) = 0$. By monotonicity of v, for any positive x_s^i, $v(x_s^i) > 0$ and $v(-x_s^i) < 0$. Therefore, the evaluation of acts is such that a higher value of probability weighting increases the value of the act when $x > 0$ and decreases the value of the bet when $x < 0$. To see how that works, consider a modified version of our running example, now with negative payoffs.

	Red ($p = 3/12$)	Yellow ($p \in [0, 6/12]$)	Black ($p \in [0, 6/12]$)	Green ($p = 3/12$)
f_{-0}	-100	0	0	-150
$f_{-\alpha}$	0	-100	0	-150

Table 11: An Ellsberg-like ambiguity example (Example 1). We also refer to the example of this table as the running example, throughout the text.

Now we have the following evaluations of each act, under our model:

$$V(f_{-0}) = \Psi(\mu(E_r))\ \omega\ v(-100) + \Psi(\mu(E_g))\ \omega_g\ v(150)$$

$$V(f_{-\alpha}) = \Psi(\mu(E_y))\ \omega\ v(-100) + \Psi(\mu(E_g))\ \omega_g\ v(-150)$$

Since $v(-150) < v(-100) < 0$, now $\Psi(\mu(E_r)) \geq \Psi(\mu(E_y))$ imply a worse evaluation of the risky act $V(f_{-0})$, reversing the result obtained in our previous example. Therefore, $\Psi(\mu(E_r)) \geq \Psi(\mu(E_y))$ would indicate ambiguity seeking behavior ($f_{-\alpha} \succeq f_{-0}$), the opposite of the ambiguity aversion ($f_0 \succeq f_\alpha$) obtained in our running example.

Moreover, as detailed in Section 5.2, our model also predicts that weighting of more ambiguous events is farther from their expected probability, but if that distortion is an over or under-weighting depends on the specific parameters of the model. Therefore, we are also able to accommodate for the stylized fact that overweighting is more typical for low expected probability events and underweighting is more typical for mid to high likelihood events, as summed up in table 10. We show that in our parametric ψ function evaluation of acts below.

$$\psi(\mu(\Pi(E_r))) = \frac{1}{4}$$

$$\psi(\mu(\Pi(E_y))) = \left(\frac{1}{4}\right)^{(1-\gamma/4)} \left(\frac{1+\eta}{4}\right)^{(\gamma/4)}$$

Again, in this specific example the preferences depend on the γ and η parameters. Specifically, for any $\gamma > 0$, $\eta \in (-1, 0) \implies f_{-\alpha} \succ f_{-0}$ and $\eta \in (0, 1) \implies f_{-0} \succ f_{-\alpha}$.

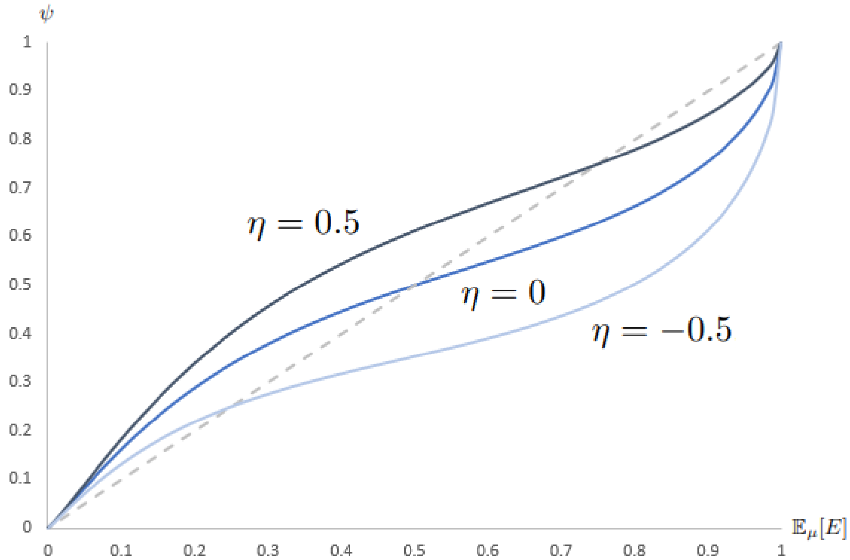
Since the ambiguity attitudes also change for different expected likelihoods of events, it is also useful to consider an additional example where we change the expected likelihoods, but nothing else. Consider the following example: let there be an urn with 100 balls, that can be either yellow or black. The DM knows that there are either 100 yellow balls (and no black balls) or 100 black balls (an no black balls). That is, $\Pi = \{\Pi_1, \Pi_2\}$, where $\Pi_1(E_y) = 1, \Pi_1(E_b) = 0$ and $\Pi_2(E_y) = 0, \Pi_2(E_b) = 1$. We analyze in the graphs below various values of $\mu'(\Pi_1(E_y)) = a$, with $a \in [0, 1]$. Note that our example is constructed so that $\mathbb{E}'_\mu(\Pi(E_y)) = \mu'(\Pi_1(E_y))$, and we assume that the partition of the state-space that is relevant is $\mathcal{P} = \{E_y, E_b\}$, so that $|S^i| = 2$ for any non-constant act f_i. In that example, then the value of our parametric $\psi(E_y)$ for an act zE_y0 with $z \neq 0$ is given by:

$$\psi(\mu(\Pi(E_y))) = (\mathbb{E}_y)^{\left(1-\gamma\cdot\sqrt{\mathbb{E}_y(1-\mathbb{E}_y)}\right)} \left(\frac{1+\eta}{2}\right)^{\left(\gamma\cdot\sqrt{\mathbb{E}_y(1-\mathbb{E}_y)}\right)}$$

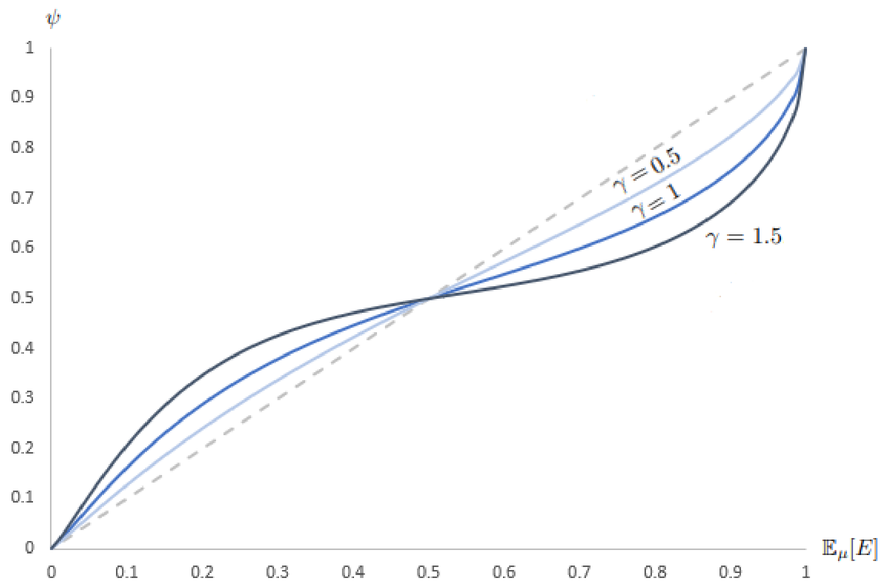
where we simplify the notation so that $\mathbb{E}_y = \mathbb{E}'_\mu(\Pi(E_y))$. We also note that, in our example, μ follows a Bernoulli distribution, and since $\mathbb{E}'_\mu(\Pi(E_y)) = \mu'(\Pi_1(E_y))$, we can express the typical $\sqrt{p(1-p)}$ standard deviation of a Bernoulli distribution as $\sqrt{\mathbb{E}_y(1-\mathbb{E}_y)}$). We show in the graph below in the x axis different values of $\mathbb{E}_y$ implied by different μ' in agreement with our example settings, while in the y axis we have values of the ψ function for different values of η and α.

In graph 3, we can see that for lower values of $\mathbb{E}_y$, typically ψ is above the 45°line, which means that there is an overweight of these low expected probability events. On the other hand, for higher expected probabilities, the ψ function has values below the 45°line, which indicate underweighting. We can interpret the point where the ψ line and the 45°line cross as the point that determines what is a "low" and what is a "high" likelihood event. The specific value that determines that inversion in the ambiguity attitude, as we described in table 10, is a question for the empirical and experimental literature, and the choice of parameters and functional forms of our model will depend on the results obtained in future research on the matter.

It is important to highlight that, in our parametric example ψ, the point that determines what is a high or low likelihood depends on how the DM responds to likelihood insensitivity, i.e., how she considers the information on the amount of states of the world in her naive probability calculations, and how these affect probability weighting. That is represented in the model by η, where lower values of η indicate that the DM needs a really low expected probability to consider it a "low likelihood" case (for which she is tipically ambiguity averse for bets on gains and ambiguity seeking for bets on losses). And, conversely, higher values of η mean that there is overweighting of events up to a higher value of $\mathbb{E}_y$. On the other hand, γ indicates the level of distortion due to the DM's sensibility to information



(a)



(b)

Figure 3: Variations in the ψ parametric function for Example 5, as a function of the expected probability of the ambiguous event E_y. Subfigure (a) shows the effect of varying η in the probability weighting function (with fixed $\gamma = 1$), while subfigure (b) shows the effect of varying γ (with fixed $\eta = 0$). Finally, the gray dashed line represents the 45°line, that is the weighting f ψ for unambiguous events.

imprecision, but not what is the indifference point that separates the region of overweighting and underweighting of probabilities. In graph 3 (b) we can see that the point where ψ crosses the 45°line remains the same for different values of γ, exemplifying rhis statement.

The careful reader will also note that these graphs have similarities with the inverse S-shaped probability weighting curves of the original Prospect Theory of Kahnemann & Tversky (1979), designed to explain behavior for choices under risk. This is not by accident: there the prediction was also an overweighting of small probability events, and underweighting of high-probability events[13]. However, the interpretation and implications here are vastly different: in our model we are saying that more imprecise information about the probability of an event (higher $\sigma(\mu(E))$) mean that the DM use information about the choice set $|S^i|$ to modify the weighting that results only from a bayesian interpretation of the available information on the expected probability of the event ($\mathbb{E}_\mu(\Pi(E))$). So, it is not only the expected probability that changes the weighting of ambiguous events, as in some adaptations of the Prospect Theory and Cumulative Prospect Theory Model (Wakker, 2010), but the choice set and the degree of information imprecision (represented by $\sigma(\mu(E))$) that determines the probability weighting for ambiguous events. Moreover, as we saw in previous examples and in Figure 1, it may be that for some events only high or only low probabilities are deemed plausible by the DM[14].

Since now we posed how the interaction between the outcome domain and expected probabilities influence the value of our Ambiguity Adjustment Function, we turn to how these same outcomes may affect salience, and that in turn may affect the results of our model.

5.4 Ambiguity Attitude, Salience and Context-Dependence

As we saw in section 3.2, the STR model proposes that salience is a characteristic that relates how the evaluation of an act by a DM is affected by the contrast of an outcome of an act f_i contingent on a state s obtaining and the outcomes associated with s that would have been obtained if other acts f_j in the choice set were chosen. In the STR model interpretation, the weighting of each state in evaluating acts is dependent on how salient the state is, where a salient state is one with highly contrasting, prominent, or surprising payoffs for the acts in the choice set. In other words, the outcomes of the choice set shape the DM's perception of the state-space (Bordalo et al., 2012).

In our extension of Salience Theory for Decision Under Ambiguity, the same principle is also present. However, its' not only the information about acts' outcomes that shape the DM's perception of the state-space, but also information on the probabilities of each event happening, how (im)precise is that information and what the partition of the choice set implies in terms of how many relevant events are deemed possible to occur. Therefore, our model does incorporate information about the choice set that is not contained in the outcome set in how the DMs evaluate acts. Again, that inclusion is crucial for models of decision under ambiguity, since the separation of an ambiguous and an unambiguous event depends on the precision of the DM's assessment of the likelihood of the event, no matter the outcomes associated. So, extending the psychological reasoning of bottom-up attention[15] influencing act evaluation through the information about the state-spaces induced by the available information of the choice set is a natural extension of the same principle to deal with decision under ambiguity.

[13] We note that more recent developments of both the original Prospect Theory and Cumulative Prospect Theory (Kahneman & Tversky, 1993) also make the probability weighting function flexible enough so that over or underweighting are possible for every probablity value, be it high or low, depending on the specific functional form and parameters chosen (Wakker, 2010). On the other hand, that flexibility also means that meaningful predictions about the DM's behavior for choice under risk are too dependent on the specific parameters chosen, which we argue can be too general a result to give meaningful predictions about economic agents' behavior, at least for Decision-Making under Ambiguity.

[14] This last consequence of our model is explored in more detail in section **??**

[15] I.e., stimuli that attract the decision-maker attention automatically and involuntarily (Bordalo et al., 2021).

We interpret the empirical evidence on likelihood insensitivity as the bottom-up attention driver of the empirical and experimental findings on ambiguity attitude (see Table 10). This is so because our model regards likelihood insensitivity as the DM's response to the information about which of the events that are relevant to the outcome of the acts in the choice set are plausible. Therefore, likelihood insensitivity is not separate from the fourfold pattern of ambiguity attitude, but part of what explains that attitude. Postulate P6 is instrumental for that to hold in the axiomatic description of our model.

Furthermore, it is only natural that we also do not assume transitivity of preferences, since that is an assumption of the STR model. That is one of the main difference between salience theory and some of the concurring approaches, such as Cumulative Prospect Theory, Choquet Expected Utility (Wakker, 2010) and other approaches based on the Expected Utility framework (Klibanoof et al., 2005; Gul e Pesendorfer, 2013). However, we do imply that, for a fixed choice set, there is consistency on how the DM weights each event. Postulate P1 makes sure there is some monotonicity to that interpretation, insofar as increasing the probability of an event unequivocally increases its' weight. Also, *ceteris paribus*, property P2 implies the DM gives less weight to events with more imprecise information. Property P5 makes sure that, for a fixed choice set, events with the same information about their probability of happening are weighted the same. More, postulate P4 makes sure that only events that are relevant to the outcomes in a choice set matter, so that the way the choice set shapes the state-space is what matters for the ambiguity adjustment function.

Another important issue that property P4 deals with is that ambiguity is a property of events (Machina, 2009) that not necessarily correspond to states in a state-space. Since the ambiguity of an event is considered to be a function of the dispersion of its' second-order probability distribution, the degree of ambiguity of an event is nonlinear, therefore calculating it separately for each state-space and calculating it directly for an event that is equivalent to the union of these states is not equivalent. In our paremetric example ψ, that is why we define $\psi_s = 1/|E|\psi_E$ for any state $s \in E \in \mathcal{P}^i$ for an act f_i. That way, we calculate the ambiguity adjustment function based on the ambiguity of the relevant event E. But, when plugging into a value function - that may imply different salience levels for states $s \in E$ - we divide the probability weighting equally among the states, so that the ambiguity level is not distorted by that. In Section 5.5 we explore that nonseparability of events issue for decision under ambiguity in greater detail.

Another question that perpasses the Salience Theory literature is whether a continuous (Bordalo et al., 2013; 2020) or a rank-based salience function (Bordalo et al., 2012) is desirable. Recently, Lanzani (2022) axiomatized STR, and argued that for representing preferences among acts with outcomes associated with correlated events, using continuous salience functions is desirable. Moreover, the continuous version of the salience function has been the most used one in the empirical literature (Dertwintel-Kalt et al., 2021; Nielsen et al., 2021). Considering that, and the fact that correlation of events' probabilities plays a highly important role in decision under ambiguity, we consider here the continuous version of salience weighting as the standard for our model.

Finally, we note that at this moment there is no conclusive evidence in the literature that indicate salience is entangled with ambiguity weighting in a way that they could be nonseparable. Therefore, we assume that salience and ambiguity weighting can be separated in our function representing the DM's preferences. Nevertheless, verifying if this kind of entaglement exists and if it is economically relevant is an important direction for future research.

5.5 Correlated Acts and Machina Reversals

Machina (2009) makes an important observation about Decision Under Ambiguity Models based on an Ellsberg-urn thought experiment: to determine the level of ambiguity of an event, and how that affects the DM's preferences, events are nonseparable. That means that - differently from what happens with the expected probability of an event, for example - the way we partition the state-space when determining preferences among acts is important and may cause preference reversals.

Here we present Machina's example: consider an urn with 101 balls, that can be either red, yellow, black or green[16]. The DM knows that 50 balls are either red or yellow, and the remaining 51 balls are either black or green. The payoffs associated with acts $f_{0,m}, f_{1,m}, f_{2,m}, f_{3,m}$ are given in the tables below:

	50 Balls		51 Balls	
	Red	Yellow	Black	Green
$f_{0,m}$	8000	8000	4000	4000
$f_{1,m}$	8000	4000	8000	4000

Table 12: Machina (2009) Ellsberg Urn first example, to illustrate how event correlation affect ambiguity perception.

	50 Balls		51 Balls	
	Red	Yellow	Black	Green
$f_{2,m}$	12000	8000	4000	0
$f_{3,m}$	12000	4000	8000	0

Table 13: Machina (2009) Ellsberg Urn second example, to illustrate how event correlation affect ambiguity perception.

$f_{0,m}$ is contingent on unambiguous acts, that is, the DM knows that she has a probability 50/101 of getting \$8000 and probability 51/101 of getting \$4000. Also, considering that she has symmetric information about how the 50/101 (51/101) probability is distributed among Red and Yellow (Black and Green) events, the expected probability of outcomes \$8000 and \$4000 should be the same for $f_{0,m}$ and $f_{1,m}$. However, since the events associated with $f_{1,m}$ outcomes are ambiguous, and considering the experimental evidence already posed in this paper, for a choice-set $F = \{f_{0,m}, f_{1,m}\}$ it should be expected that $f_{0,m} \succ f_{1,m}$.

Now consider acts $f_{2,m}$ and $f_{3,m}$. These acts are obtained by a common outcome shift from $f_{0,m}$ and $f_{1,m}$, shifting \$4000 from the Green event outcome to the Red Event outcome. If events were fully separable for decision under ambiguity, this shift should not reverse preferences, and one should expect $f_{2,m} \succ f_{3,m}$ for a choice set $F = \{f_{2,m}, f_{3,m}\}$. But this shift reverses the ambiguity properties of the upper $(f_{0,m}, f_{2,m})$ to the lower $(f_{1,m}, f_{3,m})$ acts in the choice set (Machina, 2009). Different from $f_{0,m}$, now not only $f_{2,m}$ is an ambiguous act, but the expected payoff for the more likely event that either a black or a green ball is drawn is now less than for $f_{3,m}$. Therefore, it should be within reason that the DM has $f_{3,m} \succ f_{2,m}$. The pair of preferences $f_{0,m} \succ f_{1,m}$ and $f_{3,m} \succ f_{2,m}$ is incompatible with both Subjective Expected Utility Theory and models of Rank-Dependent Preferences, such as Rank-Dependent Utility and Cumulative Prospect Theory.

[16] In Machina (2009), balls are labeled with numbers 1 through 4. That, however, is inconsequential for the results of the example.

What this example highlights is that there is a tradeoff between having more precise information about the probabilities of relevant events. $f_{0,m} \succ f_{1,m}$ is reasonable exactly because $f_{0,m}$ is contingent on unambiguous events - therefore their outcomes are dependent on events for which the DM has more precise information. That is true even though from the available information to the DM it is reasonable for her to have a higher expected payoff for $f_{1,m}$, since the group of balls with higher expected probability (black or green, with 51/101) has on average a higher average payoff. On the other hand, when the outcomes are shifted so that the advantage in terms of precision on event probability vanishes, then the shifted top lottery ($f_{2,m}$) now does not seem that much attractive, in comparison with the higher expected payoff of the bottom lottery ($f_{3,m}$).

Our model is well-suited to deal with those types of reversals. Postulate P4 makes sure that the ambiguity adjustment function reflects correctly the ambiguity of events in the Act-Induced Partition of the state-space, and that key property makes sure that the degree of ambiguity of an event is measured for events in that act-induced partition. That reflects the fact that, even though the expectation operator is linear - and therefore the expected probability of an union disjoint events can be linearly added without any issues, that is not the case for dispersion measures of a probability distribution. Therefore, it is only natural that events should be nonseparable with respect to their degree of ambiguity, and how this ambiguity level reflects in the the DM's preferences.

For concreteness, let's apply our general $V(f)$ to the Machina example. We denote the salience of each state with subscripts that have the payoffs of the acts for some state, i.e., $\omega_{(12000,12000)} < \omega_{(8000,8000)} < \omega_{(4000,4000)} < \omega_{(8000,4000)}$. First, we analyze $f_{0,m}$ and $f_{1,m}$:

$$
\begin{aligned}
V(f_{0,m}) = \frac{\Psi(\mu(E_{r \,\vee\, y})}{2}\ \omega_{(8000,8000)}\ v(8000) + \frac{\Psi(\mu(E_{r \,\vee\, y})}{2}\ \omega_{(8000,4000)}\ v(8000) \\
+ \frac{\Psi(\mu(E_{b \,\vee\, g})}{2}\ \omega_{(8000,4000)}\ v(4000) + \frac{\Psi(\mu(E_{b \,\vee\, g})}{2}\ \omega_{(4000,4000)}\ v(4000) \\
= v(8000)(\Psi(\mu(E_{r \,\vee\, y})) \cdot \omega_{(8000,8000)} + \Psi(\mu(E_{r \,\vee\, y})) \cdot \omega_{(8000,4000)}) \\
+ v(4000)(\Psi(\mu(E_{b \,\vee\, g})) \cdot \omega_{(4000,8000)} + \Psi(\mu(E_{b \,\vee\, g})) \cdot \omega_{(8000,4000)})
\end{aligned}
$$

$$
\begin{aligned}
V(f_{1,m}) = \Psi(\mu(E_r)\ \omega_{(8000,8000)}\ v(8000) + \Psi(\mu(E_y)\ \omega_{(8000,4000)}\ v(4000) \\
+ \Psi(\mu(E_b)\ \omega_{(8000,4000)}\ v(8000) + \Psi(\mu(E_g)\ \omega_{(4000,4000)}\ v(4000) \\
= v(8000)(\Psi(\mu(E_r) \cdot \omega_{(8000,8000)} + \Psi(\mu(E_b) \cdot \omega_{(8000,4000)}) + v(4000)(\Psi(\mu(E_y) \cdot \omega_{(8000,4000)} \\
+ \Psi(\mu(E_g) \cdot \omega_{(4000,4000)})
\end{aligned}
$$

Assume the DM interprets the lack of knwoledge about how much of 50 (51) balls is red and yellow (black and green) as an uniform probability of each possible combination being the true one. Then, by property P2, since $\mu(E_j|\mathcal{P}^{1,m})$ is a mean preserving spread for all events $j = r, y, b, g$, i.e., given the Act-Induced Partition of $f_{0,m}$, the relevant events for the outcomes of this act have more precise information on probability distribution and same expected probability. Therefore, $\Psi(\mu(E_{r \,\vee\, y}) \geq \Psi(\mu(E_r) + \Psi(\mu(E_y)$ and $\Psi(\mu(E_{b \,\vee\, g}) \geq \Psi(\mu(E_b) + \Psi(\mu(E_g)$. Thus, since all outcomes are positive, $f_{0,m} \succeq f_{1,m}$. Now we evaluate acts $f_{2,m}, f_{3,m}$:

$$V(f_{2,m}) = \Psi(\mu(E_r)\ \omega_{(12000,12000)}\ v(12000) + \Psi(\mu(E_y)\ \omega_{(8000,4000)}\ v(8000)$$
$$+\Psi(\mu(E_b)\ \omega_{(8000,4000)}\ v(4000) + \Psi(\mu(E_g)\ \omega_{(0,0)}\ v(0)$$
$$= \Psi(\mu(E_r)\ \omega_{(12000,12000)}\ v(12000) + \omega_{(8000,4000)}(\Psi(\mu(E_y) \cdot v(8000) + \Psi(\mu(E_b) \cdot v(4000))$$

$$V(f_{3,m}) = \Psi(\mu(E_r)\ \omega_{(12000,12000)}\ v(12000) + \Psi(\mu(E_y)\ \omega_{(8000,4000)}\ v(4000)+$$
$$\Psi(\mu(E_b)\ \omega_{(8000,4000)}\ v(8000) + \Psi(\mu(E_g)\ \omega_{(0,0)}\ v(0)$$
$$= \Psi(\mu(E_r)\ \omega_{(12000,12000)}\ v(12000) + \omega_{(8000,4000)}(\Psi(\mu(E_y) \cdot v(4000) + \Psi(\mu(E_b) \cdot v(8000))$$

Now the act-induced partitions are the same for $f_{2,m}$ and $f_{3,m}$, and the preferences depend on the relation $\Psi(\mu(E_y)$ and $\Psi(\mu(E_b)$. By the monotonicity implied by postulate P1 on second-order probability distributions, and the fact that it is clear that the expected probability of E_b is higher than E_y, $\Psi(\mu(E_b) \geq \Psi(\mu(E_y)$ and the DM is expected to have $f_{3,m} \succ f_{2,m}$, as suggested by Machina (2009).

Applying our parametric function to this example we can see again this pattern:

$$V(f_{0,m}) = v(8000)(\Psi(\mu(E_{r\ \vee\ y})) \cdot \omega_{(8000,8000)} + \Psi(\mu(E_{r\ \vee\ y})) \cdot \omega_{(8000,4000)})$$
$$+v(4000)(\Psi(\mu(E_{b\ \vee\ g})) \cdot \omega_{(4000,8000)} + \Psi(\mu(E_{b\ \vee\ g})) \cdot \omega_{(8000,4000)})$$
$$= \left(\frac{25}{101}\right)(\omega_{(8000,8000)}v(8000) + \omega_{(8000,4000)}v(8000))$$
$$+ \left(\frac{25.5}{101}\right)(\omega_{(8000,4000)}v(4000) + \omega_{(4000,4000)}v(4000))$$

$$V(f_{1,m}) = v(8000)(\Psi(\mu(E_r) \cdot \omega_{(8000,8000)} + \Psi(\mu(E_b) \cdot \omega_{(8000,4000)})$$
$$+v(4000)(\Psi(\mu(E_y) \cdot \omega_{(8000,4000)} + \Psi(\mu(E_g) \cdot \omega_{(4000,4000)})$$
$$= \left(\left(\frac{25}{101}\right)^{1-\gamma\cdot\sigma(U_d(0,50/101))}\left(\frac{1+\eta}{4}\right)^{\gamma\cdot\sigma(U_d(0,50/101))}\right)(\omega_{(8000,8000)}\ v(8000) + \omega_{(8000,4000)}\ v(4000))$$
$$+ \left(\left(\frac{25.5}{101}\right)^{1-\gamma\cdot\sigma(U_d(0,51/101))}\left(\frac{1+\eta}{4}\right)^{\gamma\cdot\sigma(U_d(0,51/101))}\right)(\omega_{(8000,4000)}\ v(8000) + \omega_{(4000,4000)}\ v(4000))$$

$$V(f_{2,m}) = \Psi(\mu(E_r)\ \omega_{(12000,12000)}\ v(12000) + \omega_{(8000,4000)}(\Psi(\mu(E_y) \cdot v(8000) + \Psi(\mu(E_b) \cdot v(4000))$$
$$= \Psi(\mu(E_r)\ \omega_{(12000,12000)}\ v(12000)$$
$$+\omega_{(8000,4000)}\left(\left(\frac{25}{101}\right)^{1-\gamma\cdot\sigma(U_d(0,50/101))}\left(\frac{1+\eta}{4}\right)^{\gamma\cdot\sigma(U_d(0,50/101))} \cdot v(8000)\right)$$
$$+\omega_{(8000,4000)}\left(\left(\frac{25.5}{101}\right)^{1-\gamma\cdot\sigma(U_d(0,51/101))}\left(\frac{1+\eta}{4}\right)^{\gamma\cdot\sigma(U_d(0,51/101))} \cdot v(4000)\right)$$

$$V(f_{3,m}) = \Psi(\mu(E_r)\ \omega_{(12000,12000)}\ v(12000) + \omega_{(8000,4000)}(\Psi(\mu(E_y)\cdot v(4000) + \Psi(\mu(E_b)\cdot v(8000))$$
$$= \Psi(\mu(E_r)\ \omega_{(12000,12000)}\ v(12000)$$
$$+\omega_{(8000,4000)}\left(\left(\frac{25}{101}\right)^{1-\gamma\cdot\sigma(U_d(0,50/101))}\left(\frac{1+\eta}{4}\right)^{\gamma\cdot\sigma(U_d(0,50/101))}\cdot v(4000)\right)$$
$$+\omega_{(8000,4000)}\left(\left(\frac{25.5}{101}\right)^{1-\gamma\cdot\sigma(U_d(0,51/101))}\left(\frac{1+\eta}{4}\right)^{\gamma\cdot\sigma(U_d(0,51/101))}\cdot v(8000)\right)$$

for any $\gamma > 0$ and $\eta \in (-1, 0)$ it is assured that $f_{0,m} \succ f_{1,m}$ and $f_{2,m} \prec f_{3,m}$, satisfying the Machina reversal.

6 Related Literature

We now briefly relate our model to other previously developed models in the literature, focusing on the most used models for decision under ambiguity and on other second-order belief models.

6.0.1 Smooth Ambiguity Preferences

One of the most popular second-order belief models for Decision Under Ambiguity is the Smooth Ambiguity Preferences model (Klibanoff et al., 2005). It avoids the problem of nondifferentiability typical of previous models, such as the α-maxmin model (Gilboa & Schmeidler, 1989). Besides, it also considers the whole first and second-order distributions - not only the most optimistic or pessimistic scenarios - in the DM's evaluation of an act. Taking Savage's Subjective Expected Utility as a starting point, and with preliminaries that are similar to our own model, in the case of a finite set of states $s \in S$ and a finite set of second-order probabilities ("scenarii") represented by μ, the value function that represents the DM's preferences is:

$$V^{SAP}(f_i) = \textstyle\sum_{\theta\in\Theta}\mu(\theta)\Phi(\pi_\theta(s)\ u(x_s))$$

In other words, the preference criterion can be read as two-layer expected utility: first, the decision-maker evaluates the expected utility with respect to all possible priors $\pi \in \Delta(S)$, so that the DM has then a set of first-order expected utilities indexed by θ. Then, the DM takes an expectation of these utilities, "distorted" by a Φ function (Etner et al., 2012). Φ, in turn, determines the ambiguity attitude of the DM, in the folllowing sense: if Φ were linear, the compound lottery representing the decision under ambiguity would just reduce to an Expected Utility problem; for concave Φ, the DM weighs more "bad" $\pi_\theta(s)u(x_s)$ in its' evaluation of results, and thus is ambiguity averse; if Φ is concave, the DM gives more weight to "good" $\pi_\theta(s)u(x_s)$, and so she is ambiguity seeking.

In that way, the decision criterion proposed by Klibanoff et al. (2005) involves both an expected utility evaluation of the possible first-order probability distributions and a pessimistic, neutral or optimistic criterion given by the Φ function. Even though the authors allow for different Φ for different supports of Π (i.e., for different sets of first-order probability distributions), when aplying the model one still has to assume a DM ambiguity attitude through the choice of the Φ function.

Particularly, even though the Smooth Ambiguity Preferences Model provides an interesting extension of the Subjective Expected Utility framework to analyze decision under ambiguity, it still does not imply any specific prediction about DM's ambiguity attitudes, nor what may influence that ambiguity attitude. Moreover, since Φ is defined on a classic SEU-like function, one cannot use the model to assume difference in probability weighting that is independent of the outcome x_s and its' associated utility function, unless if using a rather *ad hoc* approach for defining the Φ function differently for many different supports of Π. Therefore, it is hard to use this framework to predict the fourfold pattern of ambiguity attitudes empirically observed. Concretely, we can see that this usually implies that experimenters testing the Smooth Ambiguity Preferences model put an additional assumption on the Φ function to test the theory - and therefore on ambiguity attitudes (Conte & Hey, 2013; Attanasi et al., 2014; Gneezy et al., 2015).

On the other hand, our model takes advantage of the great growth in experimental and empirical evidence in recent decades to actually predict how ambiguity attitudes change as a function of the expected probability of events the outcome domain, and other specific contextual information about a given choice set. Even though assumptions about the parametric form of our model still need to be chosen and calibrated according to empirical results - as it is usual for any such model - ambiguity attitudes result from specific properties implied in our model, instead of being just assumed *ex ante* as in the Smooth Ambiguity approach. So, not only we can accommodate for different ambiguity attitudes depending on the context, we specifically predict which factors affect ambiguity attitude in a choice problem.

We also retain the interesting properties of continuity and differentiability of the act evaluation function V, for a given choice set F and beliefs μ based on the available information to the DM.

6.0.2 Choquet Expected Utility and Cumulative Prospect Theory

Rank-dependent theories, meaning models that rely on the valuation of acts according to the ranking of outcomes by the DM, have also been employed for decision under ambiguity problems. Choquet Expected Utility and Cumulative Prospect Theory, proposed by Schmeidler (1989) and Tversky & Kahneman (1992) respectively and later generalized and adapted for decision under ambiguity (Chateneuf & Faro, 2009; Chateneauf, Eichberger & Grant (2007); Wakker, 2010), are two such models.

If taken in full generality, both models can accommodate the fourfold pattern empirically observed, depending on the ambiguity parameters of the Cumulative Prospect Theory probability weighting function (Wakker, 2010). For Choquet Expected Utility, the relevance of each prior assigned in a Confidence Function such as that of Chateneuf & Fato (2009) may also accommodate those factors. However, there may be a large number of free parameters involved, so that for empirical applications a calibration of these parameters is required. Again, the models are general enough so that calibrating their free parameters may result in the fourfold pattern, but without that specific calibration we do not have *a priori* meaningful predictions about ambiguity attitude and how they change over time.

Moreover, phenomena such as Machina reversals cannot be predicted by these rank-dependent models. That is so because they both assume a form of event separability - that Machina (2009) calls tail-separability - which means that terms involving upper tail (better outcomes), mid-tail (average outcomes) and lower tail (worse outcomes) are separable, in a way that their degree of ambiguity is also separable. Therefore, tail-outcome shifts that affect ambiguity but do not affect the ranking of event-contingent outcomes cannot be accomodated by these rank-dependent models (Machina, 2009; Wakker, 2010), differently from what happens in our model.

Therefore, our model is able to deal with the nonlinearity of the degree of ambiguity based on the event-partition induced by the choice set, while that cannot be accounted for in the rank-dependent models of decision under ambiguity. This way of integrating context - interpreted as characteristics of the choice set, and not only of the specific act or outcome being evaluated - in the act evaluation is the way to account for that nonseparability of events adopted by our model.

6.0.3 Contraction Second-Order Belief Model

Gajdos et al. (2008) propose a model that contains an idea of how DM's use objective information on the probability of events that is similar to the one contained in our model. The authors give axiomatic foundations for a preference foundation that considers two criterion: (i) a Bayesian criterion, where information is summarized by the available information on probability distribution of events that is independent on the outcomes; (ii) a pessimistic criterion, so that the DM takes into account the distribution giving the lowest expected utility possible. The evaluation of an act can be represented by the function below:

$$V^{CM}(f_i) = \min_{\Phi \in \Phi^{CM}(\mu(\Pi))} \mathbb{E}_{\mu}(\Pi)\ u(f_i)$$

where $\Phi^{CM}(\mu(\Pi))$ is a subjective set of second-order priors estimated from the available information on event likelihood, and $\mathbb{E}_{\mu}(\Pi)$ is the vector of expected probabilities of each event associated with an outcome of act f_i. The Φ^{CM} function concept is similar to our Ψ transformation of the second-order subjective second-order probability distribution - the idea that objective information about ambiguous event probabilities is somehow distorted in the DM's evaluation of an act. There are, however, some important differences between our model and the Contraction Model. First, we do not assume the pessimistic criterion for the evaluation of acts, but consider that the state-space partition induced by the act and the choice set is what determines if a DM is "optimistic" or "pessimistic" about ambiguous prospects and events. In that way, ambiguity seeking behavior as a function of expected probability of events is easily accomodated by our model, while there is no clear effect of the expected probability of an ambiguous event on ambiguity attitude in the Contraction Model.

Second, we consider the whole set of priors in the DM's evaluation - not only the most pessimistic scenario. In that way, the dispersion of the second-order priors matter, and not only what is the subjective probabilities associated with the most pessimistic scenario. This kind of nuance in the DM's reaction for different degrees and forms of ambiguity is corroborated by recent experimental evidence (Chew et al., 2017).

7 Conclusion

Our article explores how contextual characteristics of decision under ambiguity may influence decision-maker's preferences. We argue that, when faced with highly imprecise information on the probabilities of each outcome (or event associated with outcomes) for different courses of action involving uncertainty, then the DM uses other information from the choice set, such as the number of possible (and relevant) events that can happen to "fill the gap" of information about the probabilities of each event with a naive equal probability distribution for each event. This bottom up stimuli distorts the weighting of different event-outcome pairs, in a phenomenon called likelihood insensitivity by the literature. Likelihood insensitivity, in turn, causes the DM to overweight (underweight) low (high) likelihood ambiguous events, causing the fourfold ambiguity attitude observed in the literature.

However, our postulates also imply some ways in which decision-maker's are consistent when dealing with ambiguous events. Property P1 makes sure that, for a given expected probability of an event, any monotonic increase in probability increases the weighting of an event. For example, an event that has between 10% and 11% of happening is going to be weighted more than an unambiguous event with

10% chance of happening. Property P2 makes sure that people weight coherently events based on the (im)precision of available information on their probability, for a given value of expected probability. That is, for a given level of expected probability, adding noise to second order probabilities (i.e., making an event more ambiguous) either monotonically increases or decreases the weighting of the event. Postulate P3 makes sure probability weighting functions are continuous. Postulate P4 assures that only the ambiguity of events that are relevant to the acts' outcomes matter in event weighting, while P5 states that events weighting depend exclusively on the available information on probabilities and information on the amount of possible events that can affect the outcomes of acts. Finally, P6 introduces likelihood insensitivity as a function of this amount of relevant events (the Act-Induced Partition). It introduces how a naive equal probability of each relevant event is used by the DM to classify an event as a low or high-likelihood, and then to adjust their weighting through likelihood insensitivity, to generate the fourfold pattern of Ambiguity Attitudes.

This model specification contrast with previous ones as we define postulates not about preferences over acts themselves, but about the DM's interpretation of available information on outcome/event probabilities and the context of the decision, given by choice set information.

For future research, some interesting questions arise: (i) is there a limit to how many events-outcomes can be considered by a decision-maker when calibrating his weighting of events with a naive probability distribution: That is, if there are 100 events relevant to an act's outcome, is the cutoff to define a "low" expected probability of an event less than the cutoff when there are 99 relevant events? Or is there a limit to this cutoff point? Is there an interaction between sensitivity to salience and sensitivity to distortions in probability weighting? In other words, are people who are more affected by bottom up salience are also more affected by bottom up Contextual-Ambiguity distortions in event weighting? These questions will certainly provide great insights to calibrate and apply the proposed model many different puzzles related to decision under ambiguity.

Appendix A: Proofs of Propositions

The $\psi(\mu(\Pi(s)))$ is defined as follows:

$$\psi(\mu(\Pi(s))) = \frac{1}{|E_s^i|}\left(\mathbb{E}_\mu(\Pi(E_s^i))\right)^{1-\gamma\cdot\sigma_\mu(\Pi(E_s^i))}\left(\frac{1+\eta}{|S^i|}\right)^{\gamma\cdot\sigma_\mu(\Pi(E_s^i))} \& \eta \in (-1,1),\ \gamma \in (0,2) \tag{2}$$

Lemma 1: Assume Nontriviality (A2). Then, the ψ function is: (i) Well-defined. (ii) Continuous on $\mu(\Pi(s))$

Proof:

To prove (i), first, note that, the marginal probability $\Pi(E_s^i) \in [0,1]$ is a bounded random variable. Then, assuming nontriviality - so that a distribution μ can be well-defined on $\Pi(E_s^i)$, by the bounds on $\Pi(E_s^i)$, $\mu(\Pi(E_s^i))$ has finite variance and expectation. Therefore, $\mathbb{E}_\mu(\Pi(E_s^i)$ and $\sigma_\mu(\Pi(E_s^i))$ are well-defined for every $s \in E_s^i \in \mathcal{P}^i$. $|E_s^i|$ is the cardinality of the event $E_s^i \in \mathcal{P}^i$ that contains state s. By the fact that an Act-Induced Partition is a partition of a non-empty finite and discrete S, E_s^i is non-empty and finite. Therefore, $|E_s^i| \in \mathbb{N} : |E_s^i| > 1$ is constant for a given act f_i and state s. Similarly, by the definition of partition of a set, since S^i is a partition of S, S^i is finite and non-empty and $|S^i| \in \mathbb{N} : |S^i| \geq 1$. Therefore, every term in the function is well defined in the real numbers, and thus ψ is well-defined.

To prove (ii), see that since $\mathbb{E}_\mu(\Pi(E))$ and $\sigma_\mu(\Pi(E))$ are sums and products of continuous functions (i.e., probabilities), they are continuous functions of $\mu(\Pi(s))$. Thus, $\psi(\mu(\Pi(s)))$ a continuous function of $\mu(\Pi(s))$.

Lemma 2: Let $\Pi(E_s^i) \in [0,1]$ be a discrete bounded random variable, $\mathcal{A}(\Pi(E_s^i))$ its' σ-algebra and μ a probability distribution measure defined on the measurable space $(\Pi(E_s^i), \mathcal{A}(\Pi(E_s^i)))$. Then, $\max \sigma_\mu(\Pi(E_s^i)) = \frac{1}{2}$.

Proof:

Note that $\sigma_\mu(\Pi(E_s^i))$ is a measure of dispersion of $\Pi(E_s^i)$ under the μ distribution. Since $\Pi(E_s^i)$ is bounded, it the maximum dispersion is achieved when the probability mass is concentrated in its' extremes, i.e., when $\Pi(E_s^i) = 0$ and $\Pi(E_s^i) = 1$, each with probability 1/2. In this case, $\max \sigma_\mu(\Pi(E_s^i)) = \sqrt{\frac{1}{2}(0-1/2)^2 + \frac{1}{2}(1-1/2)^2} = \frac{1}{2}$.

ψ satisfies property P1

Proof:

By the Lemma, ψ is well-defined and continuous on its' arguments.

First, define:

$$g(\mu(\Pi(s))) = \mathbb{E}_\mu(\Pi(s))^{1-\gamma\cdot\sigma_\mu(\Pi(s))} \qquad\qquad h(\mu(\Pi(s))) = \left(\frac{1+\eta}{|\mathcal{P}^i|}\right)^{\gamma\cdot\sigma_\mu(\Pi(s))}$$

Then, by the product rule:

$$\frac{\partial\psi(\mu(\Pi(E)))}{\partial\mu(\Pi(s))} = \frac{\partial g(\mu(\Pi(s)))}{\partial\mu(\Pi(s))} \cdot h(\mu(\Pi(s))) + g(\mu(\Pi(s))) \cdot \frac{\partial h(\mu(\Pi(s)))}{\partial\mu(\Pi(s))}$$

Applying the chain rule to g:

$$\frac{\partial g(\mu(\Pi(s)))}{\partial\mu(\Pi(s))} = \frac{\partial g}{\partial\mathbb{E}_\mu(\Pi(s))} \cdot \frac{\partial\mathbb{E}_\mu(\Pi(s))}{\partial\mu(\Pi(s))} + \frac{\partial g}{\partial\sigma_\mu(\Pi(s))} \cdot \frac{\partial\sigma_\mu(\Pi(s))}{\partial\mu(\Pi(s))} \tag{3}$$

And the chain rule applied to h gives:

$$\frac{\partial h(\mu(\Pi(s)))}{\partial\mu(\Pi(s))} = \frac{\partial h}{\partial\sigma_\mu(\Pi(s))} \cdot \frac{\partial\sigma_\mu(\Pi(s))}{\partial\mu(\Pi(s))} \tag{4}$$

Now, we need to compute the partial derivatives of g and h with respect to $\mathbb{E}_\mu(\Pi(s))$ and $\sigma_\mu(\Pi(s))$, and the partial derivatives of $\mathbb{E}_\mu(\Pi(s))$ and $\sigma_\mu(\Pi(s))$ with respect to $\mu(\Pi(s))$. After computing these derivatives and substituting them back into the expressions for $\frac{\partial g(\mu(\Pi(s)))}{\partial\mu(\Pi(s))}$ and $\frac{\partial h(\mu(\Pi(s)))}{\partial\mu(\Pi(s))}$, we can simplify the resulting expression for $\frac{\partial\psi(\mu(\Pi(s)))}{\partial\mu(\Pi(s))}$.

To compute the partial derivatives of g and h with respect to $\mathbb{E}_\mu(\Pi(s))$ and $\sigma_\mu(\Pi(s))$, we first differentiate g with respect to $\mathbb{E}_\mu(\Pi(s))$:

$$\frac{\partial g}{\partial \mathbb{E}_\mu(\Pi(s))} = (1 - \gamma \cdot \sigma_\mu(\Pi(s)))\mathbb{E}_\mu(\Pi(s))^{-\gamma \cdot \sigma_\mu(\Pi(s))}$$

Next, we differentiate g with respect to $\sigma_\mu(\Pi(s))$:

$$\frac{\partial g}{\partial \sigma_\mu(\Pi(s))} = -\gamma \cdot \mathbb{E}_\mu(\Pi(s))^{1-\gamma \cdot \sigma_\mu(\Pi(s))} \cdot \log(\mathbb{E}_\mu(\Pi(s)))$$

Now, we differentiate h with respect to $\sigma_\mu(\Pi(s))$:

$$\frac{\partial h}{\partial \sigma_\mu(\Pi(s))} = \gamma \cdot \left(\frac{1+\eta}{|\mathcal{P}^i|}\right)^{\gamma \cdot \sigma_\mu(\Pi(s))} \cdot \log\left(\frac{1+\eta}{|\mathcal{P}^i|}\right)$$

Next, we compute the partial derivatives of $\mathbb{E}_\mu(\Pi(s))$ and $\sigma_\mu(\Pi(s))$ with respect to $\mu(\Pi(s))$.

For $\mathbb{E}_\mu(\Pi(s))$:

$$\frac{\partial \mathbb{E}_\mu(\Pi(s))}{\partial \mu(\Pi(s))} = \pi_i(s)$$

For $\sigma_\mu(\Pi(s))$:

$$\frac{\partial \sigma_\mu(\Pi(s))}{\partial \mu(\Pi(s))} = \frac{1}{2\sqrt{\sum_{i=1}^N p(\pi_i)(\pi_i(s) - \mathbb{E}_\mu(\Pi(s)))}} \cdot (\pi_i(s) - \mathbb{E}_\mu(\Pi(s)))$$

Substituting these derivatives back into the expressions for $\frac{\partial g(\mu(\Pi(s)))}{\partial \mu(\Pi(s))}$ and $\frac{\partial h(\mu(\Pi(s)))}{\partial \mu(\Pi(s))}$ and simplifying, we get:

$$\frac{\partial \psi(\mu(\Pi(s)))}{\partial \mu(\Pi(s))} = (\pi_i(s) - \gamma \cdot \log(\mathbb{E}_\mu(\Pi(s))) \cdot (\pi_i(s) - \mathbb{E}_\mu(\Pi(s)))) \cdot g(\mu(\Pi(s))) \cdot h(\mu(\Pi(s)))$$

The final step is to show that this expression for $\frac{\partial \psi(\mu(\Pi(s)))}{\partial \mu(\Pi(s))}$ is positive. Since $\gamma \in [0, 2)$ and $\eta \in (-1, 1)$, the terms involved in the expression are all positive or non-negative. Therefore, the expression for the derivative will be positive. As $\gamma \in [0, 2)$, $\eta \in (-1, 1)$, and $\pi_i(s) \in [0, 1]$, the terms involved in the expression are all positive or non-negative. Thus, the expression for the derivative, $\frac{\partial \psi(\mu(\Pi(s(s)))}{\partial \mu(\Pi(s))}$, is positive.

Step 2: Define the probability distributions involved in the property (P1), and show how λ affects the compound probability distribution.

Let $\mu_1(\Pi(s))$, $\mu_2(\Pi(s))$, and $\mu_i(\Pi(s))$ be the compound probability distributions defined in the property (P1). We can define the following probability distributions:

$$\mu_1'(\Pi(s)) = \lambda\mu_1(\Pi(s)) + (1-\lambda)\mu_i(\Pi(s)) \; \mu_2'(\Pi(s)) \qquad = \lambda\mu_2(\Pi(s)) + (1-\lambda)\mu_i(\Pi(s))$$

Step 3: Prove the property (P1) using the results from steps 1 and 2.

We want to show that:

$$\psi(\mu_1(\Pi(s))) \geq \psi(\mu_2(\Pi(s))) \Leftrightarrow \psi(\mu_1'(\Pi(s))) \geq \psi(\mu_2'(\Pi(s)))$$

We have shown in step 1 that $\psi(\mu(\Pi(s)))$ is a continuous and increasing function of $\mu(\Pi(s))$. Since $\mu_1(\Pi(s)) \geq \mu_2(\Pi(s))$, we can apply the continuous and increasing property of the function ψ:

$$\psi(\mu_1(\Pi(s))) \geq \psi(\mu_2(\Pi(s))) \Leftrightarrow \psi(\mu_1'(\Pi(s))) \geq \psi(\mu_2'(\Pi(s)))$$

Now, consider the compound probability distributions $\mu_1'(\Pi(s))$ and $\mu_2'(\Pi(s))$:

$$\mu_1'(\Pi(s)) = \lambda\mu_1(\Pi(s)) + (1-\lambda)\mu_i(\Pi(s)) \; \mu_2'(\Pi(s)) \qquad = \lambda\mu_2(\Pi(s)) + (1-\lambda)\mu_i(\Pi(s))$$

Since $\lambda \in (0, 1)$, the compound probability distributions are convex combinations of $\mu_1(\Pi(s))$, $\mu_2(\Pi(s))$, and $\mu_i(\Pi(s))$. Therefore, the compound probability distributions are also continuous and increasing functions of $\mu(\Pi(s))$. Moreover, due to the convexity of the combination, we have that:

$$\mu_1'(\Pi(s)) \geq \mu_2'(\Pi(s)) \Leftrightarrow \lambda\mu_1(\Pi(s)) + (1-\lambda)\mu_i(\Pi(s)) \geq \lambda\mu_2(\Pi(s)) + (1-\lambda)\mu_i(\Pi(s))$$

Applying the continuous and increasing property of the function ψ to the compound probability distributions, we have:

$$\psi(\mu_1'(\Pi(s))) \geq \psi(\mu_2'(\Pi(s))) \Leftrightarrow \psi(\lambda\mu_1(\Pi(s)) + (1-\lambda)\mu_i(\Pi(s))) \geq \psi(\lambda\mu_2(\Pi(s)) + (1-\lambda)\mu_i(\Pi(s)))$$

ψ satisfies property P2

We will prove the property (P2) for the given function $\psi(\mu(\Pi(s)))$ in the following steps:

Show that the function $\psi(\mu(\Pi(s)))$ is continuous and differentiable with respect to $\mu(\Pi(s))$. Derive expressions for the partial derivatives $\frac{\partial \sigma_\mu(\Pi(s))}{\partial \mu(\Pi(s))}$ and $\frac{\partial \mathbb{E}_\mu(\Pi(s))}{\partial \mu(\Pi(s))}$. Analyze how the partial derivatives affect the inequality conditions. Show that the conditions for property (P2) hold. Step 1: Continuity and differentiability of $\psi(\mu(\Pi(s)))$

The given function $\psi(\mu(\Pi(s)))$ is a composition of continuous functions: power functions, products, and sums. Therefore, $\psi(\mu(\Pi(s)))$ is continuous. Furthermore, since all the functions involved are differentiable, $\psi(\mu(\Pi(s)))$ is also differentiable.

Step 2: Deriving expressions for partial derivatives

We first compute the partial derivatives of $\mathbb{E}_\mu(\Pi(s))$ and $\sigma_\mu(\Pi(s))$ with respect to $\mu(\Pi(s))$.

For $\mathbb{E}_\mu(\Pi(s))$, we have:

$$\frac{\partial \mathbb{E}\mu(\Pi(s))}{\partial \mu(\Pi(s))} = \frac{\partial}{\partial \mu(\Pi(s))} \left(\sum^N i = 1 p(\pi_i) \times \pi_i(s) \right) = \pi_i(s)$$

For $\sigma_\mu(\Pi(s))$, we have:

$$\frac{\partial \sigma_\mu(\Pi(s))}{\partial \mu(\Pi(s))} = \frac{\partial}{\partial \mu(\Pi(s))} \sqrt{\sum_{i=1}^{N} p(\pi_i)(\pi_i(s) - \mathbb{E}_\mu(\Pi(s)))^2}$$

We will not explicitly compute the partial derivative of $\sigma_\mu(\Pi(s))$ with respect to $\mu(\Pi(s))$, as it is not necessary for our analysis.

Step 3: Analyzing the effect of partial derivatives on the inequality conditions

Since $\psi(\mu(\Pi(s)))$ is continuous and differentiable, the conditions in property (P2) can be analyzed using the partial derivatives. We need to show that the following inequality holds:

$$\psi(\lambda\ \mu_1(E) + (1-\lambda)\ \mu_3(E)) \geq \psi(\lambda'\ \mu_1(E) + (1-\lambda')\ \mu_3(E))$$

if and only if

$$\psi(\lambda\ \mu_2(E) + (1-\lambda)\ \mu_4(E)) \geq \psi(\lambda'\ \mu_2(E) + (1-\lambda')\ \mu_4(E))$$

Step 4: Showing the conditions for property (P2) hold

We will use the given function definition and the partial derivative expressions to analyze the inequalities.

Since $\gamma \in [0, 2)$, the exponent $\gamma \cdot \sigma_\mu(\Pi(s))$ is non-negative. Therefore, the function $\psi(\mu(\Pi(s)))$ is non-decreasing with respect to $\sigma_\mu(\Pi(s))$. Additionally, since $\eta \in (-1, 1)$, the expression $(\frac{1+\eta}{|\mathcal{P}^i|})^{\gamma \cdot \sigma_\mu(\Pi(s))}$ is always positive. Therefore, the sign of the inequality will depend on the term $\mathbb{E}_\mu(\Pi(s))^{1-\gamma \cdot \sigma_\mu(\Pi(s))}$.

Consider the inequality:

$$\psi(\lambda\ \mu_1(E) + (1-\lambda)\ \mu_3(E)) \geq \psi(\lambda'\ \mu_1(E) + (1-\lambda')\ \mu_3(E))$$

Using the definition of $\psi(\mu(\Pi(s)))$, this inequality can be expressed as:

$$\mathbb{E}\lambda\mu_1 + (1-\lambda)\mu_3(\Pi(s))^{1-\gamma \cdot \sigma\lambda\mu_1 + (1-\lambda)\mu_3(\Pi(s))} \geq \mathbb{E}\lambda'\mu_1 + (1-\lambda')\mu_3(\Pi(s))^{1-\gamma \cdot \sigma\lambda'\mu_1 + (1-\lambda')\mu_3(\Pi(s))}$$

Now, consider the inequality:

$$\psi(\lambda\ \mu_2(E) + (1-\lambda)\ \mu_4(E)) \geq \psi(\lambda'\ \mu_2(E) + (1-\lambda')\ \mu_4(E))$$

Similarly, using the definition of $\psi(\mu(\Pi(s)))$, this inequality can be expressed as:

$$\mathbb{E}\lambda\mu_2 + (1-\lambda)\mu_4(\Pi(s))^{1-\gamma \cdot \sigma\lambda\mu_2 + (1-\lambda)\mu_4(\Pi(s))} \geq \mathbb{E}\lambda'\mu_2 + (1-\lambda')\mu_4(\Pi(s))^{1-\gamma \cdot \sigma\lambda'\mu_2 + (1-\lambda')\mu_4(\Pi(s))}$$

Given that $\psi(\mu_1(E)) > \psi(\mu_2(E))$ and $\psi(\mu_3(E)) > \psi(\mu_4(E))$, we can infer that:

$$\mathbb{E}\mu_1(\Pi(s))^{1-\gamma\cdot\sigma\mu_1(\Pi(s))} > \mathbb{E}\mu_2(\Pi(s))^{1-\gamma\cdot\sigma\mu_2(\Pi(s))}$$

and

$$\mathbb{E}\mu_3(\Pi(s))^{1-\gamma\cdot\sigma\mu_3(\Pi(s))} > \mathbb{E}\mu_4(\Pi(s))^{1-\gamma\cdot\sigma\mu_4(\Pi(s))}$$

ψ satisfies property P3

To prove Property (P3), we will proceed in the following steps:

State and prove a lemma that describes the relationship between the partial derivatives of $\sigma_\mu(\Pi(s))$ and $\mathbb{E}_\mu(\Pi(s))$ with respect to $\mu(\Pi(s))$. Analyze the impact of these partial derivatives on the function $\psi(\mu(\Pi(s)))$. Demonstrate that the function satisfies Property (P3) using the findings from steps 1 and 2. Step 1: Lemma and its proof

Lemma: For any $s \in S$, the partial derivatives of $\sigma_\mu(\Pi(s))$ and $\mathbb{E}_\mu(\Pi(s))$ with respect to $\mu(\Pi(s))$ are as follows:

(i) $\frac{\partial \mathbb{E}_\mu(\Pi(s))}{\partial \mu(\Pi(s))} = \pi_i(s) - \mathbb{E}_\mu(\Pi(s))$. (ii) $\frac{\partial \sigma_\mu(\Pi(s))}{\partial \mu(\Pi(s))} = (\pi_i(s) - \mathbb{E}_\mu(\Pi(s)))^2 - \sigma_\mu^2(\Pi(s))$.

Proof:

(i) Using the definition of $\mathbb{E}_\mu(\Pi(s))$, we have:

$\frac{\partial \mathbb{E}\mu(\Pi(s))}{\partial \mu(\Pi(s))} = \frac{\partial}{\partial \mu(\Pi(s))} \sum^N i = 1 p(\pi_i) \times \pi_i(s)$.

Since the only term involving $\mu(\Pi(s))$ is $p(\pi_i) \times \pi_i(s)$, the derivative is:

$\frac{\partial \mathbb{E}_\mu(\Pi(s))}{\partial \mu(\Pi(s))} = \pi_i(s) - \mathbb{E}_\mu(\Pi(s))$.

(ii) Using the definition of $\sigma_\mu(\Pi(s))$, we have:

$\frac{\partial \sigma_\mu(\Pi(s))}{\partial \mu(\Pi(s))} = \frac{\partial}{\partial \mu(\Pi(s))} \sqrt{\sum_{i=1}^N p(\pi_i)(\pi_i(s) - \mathbb{E}_\mu(\Pi(s)))^2}$.

We can rewrite $\sigma_\mu^2(\Pi(s))$ as:

$\sigma_\mu^2(\Pi(s)) = \sum_{i=1}^N p(\pi_i)(\pi_i(s) - \mathbb{E}_\mu(\Pi(s)))^2$.

Taking the derivative with respect to $\mu(\Pi(s))$, we get:

$\frac{\partial \sigma_\mu^2(\Pi(s))}{\partial \mu(\Pi(s))} = (\pi_i(s) - \mathbb{E}_\mu(\Pi(s)))^2 - \sigma_\mu^2(\Pi(s))$.

Step 2: Analyze the impact of partial derivatives on $\psi(\mu(\Pi(s)))$

From the given function definition, we can observe that:

$\mathbb{E}_\mu(\Pi(s))$ is increasing with respect to $\mu(\Pi(s))$. $\sigma_\mu(\Pi(s))$ is a non-decreasing function of $\mu(\Pi(s))$.

Step 3: Demonstrate Property (P3)

Let μ_i, μ_j, μ_k be such that $\psi(\mu_i(\Pi(E))) > \psi(\mu_j(\Pi(E)))$. We need to show that there exists a $\lambda \in (0,1)$ such that:

$\psi(\mu_i(\Pi(E))) > \psi(\lambda\, \mu_k(\Pi(E)) + (1-\lambda)\, \mu_j(\Pi(E)))$. $\psi(\lambda\, \mu_k(\Pi(E)) + (1-\lambda)\, \mu_i(\Pi(E))) > \psi(\mu_j(\Pi(E)))$. Consider the function $\psi(\mu(\Pi(E)))$ and its partial derivatives. We know that $\frac{\partial \mathbb{E}_\mu(\Pi(E))}{\partial \mu(\Pi(E))} = \pi_i(E) - \mathbb{E}_\mu(\Pi(E))$ and $\frac{\partial \sigma_\mu(\Pi(E))}{\partial \mu(\Pi(E))} = (\pi_i(E) - \mathbb{E}_\mu(\Pi(E)))^2 - \sigma_\mu^2(\Pi(E))$.

Now, we want to analyze the impact of these partial derivatives on the function $\psi(\mu(\Pi(E)))$. We can write the function as:

$$\psi(\mu(\Pi(E))) = \mathbb{E}_\mu(\Pi(E))^{1-\gamma\cdot\sigma_\mu(\Pi(E))} \left(\frac{1+\eta}{|\mathcal{P}^i|}\right)^{\gamma\cdot\sigma_\mu(\Pi(E))}$$

Taking the derivative of ψ with respect to $\mu(\Pi(E))$, we have:

$$\frac{\partial \psi(\mu(\Pi(E)))}{\partial \mu(\Pi(E))} = \left(\frac{1-\gamma \cdot \sigma_\mu(\Pi(E))}{\mathbb{E}_\mu(\Pi(E))} \cdot \frac{\partial \mathbb{E}_\mu(\Pi(E))}{\partial \mu(\Pi(E))} + \frac{\gamma}{|\mathcal{P}^i|} \cdot \frac{\partial \sigma_\mu(\Pi(E))}{\partial \mu(\Pi(E))} \right) \cdot \psi(\mu(\Pi(E)))$$

Since $\psi(\mu(\Pi(E))) > 0$, the sign of the derivative depends on the term inside the parentheses.

Now, we know that $\mathbb{E}_\mu(\Pi(E))$ is increasing with respect to $\mu(\Pi(E))$, and $\sigma_\mu(\Pi(E))$ is a non-decreasing function of $\mu(\Pi(E))$. Thus, for any $0 < \lambda < 1$, we have:

$\mathbb{E}\lambda\ \mu_k(\Pi(E)) + (1-\lambda)\ \mu_i(\Pi(E))(\Pi(E)) < \mathbb{E}\mu_i(\Pi(E))$ and $\mathbb{E}\lambda\ \mu_k(\Pi(E)) + (1-\lambda)\ \mu_j(\Pi(E))(\Pi(E)) > \mathbb{E}\mu_j(\Pi(E))$. $\sigma_{\lambda\ \mu_k(\Pi(E))+(1-\lambda)\ \mu_i(\Pi(E))}(\Pi(E)) < \sigma_{\mu_i}(\Pi(E))$ and $\sigma_{\lambda\ \mu_k(\Pi(E))+(1-\lambda)\ \mu_j(\Pi(E))}(\Pi(E)) > \sigma_{\mu_j}(\Pi(E))$.

Since $\sigma_\mu(\Pi(E))$ is non-decreasing with respect to $\mu(\Pi(E))$, we can find $\lambda \in (0,1)$ such that the term inside the parentheses of the derivative expression is positive for $\lambda\ \mu_k(\Pi(E)) + (1-\lambda)\ \mu_i(\Pi(E))$ and negative for $\lambda\ \mu_k(\Pi(E)) + (1-\lambda)\ \mu_j(\Pi(E))$. This implies that the derivative of ψ with respect to $\mu(\Pi(E))$ is positive for $\lambda\ \mu_k(\Pi(E)) + (1-\lambda)\ \mu_i(\Pi(E))$ and negative for $\lambda\ \mu_k(\Pi(E)) + (1-\lambda)\ \mu_j(\Pi(E))$.

Therefore, we have:

$\psi(\mu_i(\Pi(E))) > \psi(\lambda\ \mu_k(\Pi(E)) + (1-\lambda)\ \mu_j(\Pi(E)))$ since the derivative of ψ is negative for $\lambda\ \mu_k(\Pi(E)) + (1-\lambda)\ \mu_j(\Pi(E))$. $\psi(\lambda\ \mu_k(\Pi(E)) + (1-\lambda)\ \mu_i(\Pi(E))) > \psi(\mu_j(\Pi(E)))$ since the derivative of ψ is positive for $\lambda\ \mu_k(\Pi(E)) + (1-\lambda)\ \mu_i(\Pi(E))$. Thus, we have shown that there exists a $\lambda \in (0,1)$ satisfying the conditions of Property (P3).

ψ satisfies property P4

To prove property (P4) for the given function $\psi(\mu(\Pi(s)))$, we will proceed with the following steps:

Prove that $\psi(\mu(\Pi(s)))$ is strictly increasing in $\mathbb{E}_\mu(\Pi(s))$ and strictly decreasing in $\sigma_\mu(\Pi(s))$. Show that the convex combination of probability measures, $\lambda\ \mu_k(\Pi(s)) + (1-\lambda)\ \mu_j(\Pi(s))$, leads to an increase in $\mathbb{E}_\mu(\Pi(s))$ and a decrease in $\sigma_\mu(\Pi(s))$. Apply the results of steps 1 and 2 to verify property (P4). Step 1: Prove that $\psi(\mu(\Pi(s)))$ is strictly increasing in $\mathbb{E}_\mu(\Pi(s))$ and strictly decreasing in $\sigma_\mu(\Pi(s))$.

First, we take the partial derivative of $\psi(\mu(\Pi(s)))$ with respect to $\mathbb{E}_\mu(\Pi(s))$:

$$\frac{\partial \psi(\mu(\Pi(s)))}{\partial \mathbb{E}_\mu(\Pi(s))} = (1-\gamma \cdot \sigma_\mu(\Pi(s)))\mathbb{E}_\mu(\Pi(s))^{-\gamma \cdot \sigma_\mu(\Pi(s))} \left(\frac{1+\eta}{|\mathcal{P}^i|} \right)^{\gamma \cdot \sigma_\mu(\Pi(s))}$$

Since $\gamma \in [0,2)$, and $\sigma_\mu(\Pi(s))$ is non-negative, the term $(1-\gamma \cdot \sigma_\mu(\Pi(s)))$ is positive. Moreover, since $\mathbb{E}_\mu(\Pi(s))$ and the term inside the parentheses are also positive, the partial derivative is positive. Thus, $\psi(\mu(\Pi(s)))$ is strictly increasing in $\mathbb{E}_\mu(\Pi(s))$.

Next, we take the partial derivative of $\psi(\mu(\Pi(s)))$ with respect to $\sigma_\mu(\Pi(s))$:

$$\frac{\partial \psi(\mu(\Pi(s)))}{\partial \sigma_\mu(\Pi(s))} = -\gamma \cdot \mathbb{E}_\mu(\Pi(s))^{1-\gamma \cdot \sigma_\mu(\Pi(s))} \left(\frac{1+\eta}{|\mathcal{P}^i|} \right)^{\gamma \cdot \sigma_\mu(\Pi(s))} \left(\ln(\mathbb{E}_\mu(\Pi(s))) - \ln\left(\frac{1+\eta}{|\mathcal{P}^i|} \right) \right)$$

Since $\gamma \geq 0$, and all the other terms inside the parentheses are positive, the partial derivative is negative. Thus, $\psi(\mu(\Pi(s)))$ is strictly decreasing in $\sigma_\mu(\Pi(s))$.

Step 2: Show that the convex combination of probability measures, $\lambda\ \mu_k(\Pi(s)) + (1-\lambda)\ \mu_j(\Pi(s))$, leads to an increase in $\mathbb{E}_\mu(\Pi(s))$ and a decrease in $\sigma_\mu(\Pi(s))$.

Let $\mu_i = \lambda\ \mu_k + (1-\lambda)\ \mu_j$. Then,

$$\mathbb{E}\mu_i(\Pi(s)) = \sum i = 1^N p(\pi_i) \times \pi_i(s) = \lambda \sum_{i=1}^{N} p_k(\pi_i) \times \pi_i(s) + (1-\lambda) \sum_{i=1}^{N} p_j(\pi_i) \times \pi_i(s)$$

Since $\psi(\mu_i(\Pi(s))) > \psi(\mu_j(\Pi(s)))$, it follows that $\mathbb{E}\mu_i(\Pi(s)) > \mathbb{E}\mu_j(\Pi(s))$. Thus, the convex combination of probability measures leads to an increase in $\mathbb{E}_\mu(\Pi(s))$.

To show that the convex combination leads to a decrease in $\sigma_\mu(\Pi(s))$, we first note that the variance of a convex combination of probability measures is:

$$\sigma^2_{\mu_i}(\Pi(s)) = \lambda^2 \sigma^2_{\mu_k}(\Pi(s)) + (1-\lambda)^2 \sigma^2_{\mu_j}(\Pi(s)) + 2\lambda(1-\lambda)\text{Cov}_{\mu_k,\mu_j}(\Pi(s))$$

Since $\sigma^2_{\mu_k}(\Pi(s))$ and $\sigma^2_{\mu_j}(\Pi(s))$ are non-negative, and by the Cauchy-Schwarz inequality, $\text{Cov}\mu_k, \mu_j(\Pi(s))^2 \leq \sigma^2 \mu_k(\Pi(s))\sigma^2_{\mu_j}(\Pi(s))$, we have:

$$\sigma^2_{\mu_i}(\Pi(s)) \leq \lambda^2 \sigma^2_{\mu_k}(\Pi(s)) + (1-\lambda)^2 \sigma^2_{\mu_j}(\Pi(s))$$

Thus, the convex combination of probability measures leads to a decrease in $\sigma_\mu(\Pi(s))$.

Step 3: Apply the results of steps 1 and 2 to verify property (P4).

From Step 1, we know that $\psi(\mu(\Pi(s)))$ is strictly increasing in $\mathbb{E}_\mu(\Pi(s))$ and strictly decreasing in $\sigma_\mu(\Pi(s))$. From Step 2, we showed that the convex combination of probability measures, $\lambda\ \mu_k(\Pi(s)) + (1-\lambda)\ \mu_j(\Pi(s))$, leads to an increase in $\mathbb{E}_\mu(\Pi(s))$ and a decrease in $\sigma_\mu(\Pi(s))$.

Now, let's consider $\psi(\lambda\ \mu_k(\Pi(s)) + (1-\lambda)\ \mu_j(\Pi(s)))$. Since the convex combination increases $\mathbb{E}_\mu(\Pi(s))$ and decreases $\sigma_\mu(\Pi(s))$, it follows that:

$$\psi(\mu_i(\Pi(s))) > \psi(\lambda\ \mu_k(\Pi(s)) + (1-\lambda)\ \mu_j(\Pi(s)))$$

Similarly, for $\psi(\lambda\ \mu_k(\Pi(s)) + (1-\lambda)\ \mu_i(\Pi(s)))$, the convex combination also increases $\mathbb{E}_\mu(\Pi(s))$ and decreases $\sigma_\mu(\Pi(s))$, which implies:

$$\psi(\lambda\ \mu_k(\Pi(s)) + (1-\lambda)\ \mu_i(\Pi(s))) > \psi(\mu_j(\Pi(s)))$$

Thus, property (P4) holds for the given function $\psi(\mu(\Pi(s)))$. We have shown that for all μ_i, μ_j, μ_k satisfying $\psi(\mu_i(\Pi(s))) > \psi(\mu_j(\Pi(s)))$, there exists $\lambda \in (0,1)$ such that:

$$\psi(\mu_i(\Pi(s))) > \psi(\lambda\ \mu_k(\Pi(s)) + (1-\lambda)\ \mu_j(\Pi(s))) \text{ and } \psi(\lambda\ \mu_k(\Pi(s)) + (1-\lambda)\ \mu_i(\Pi(s))) > \psi(\mu_j(\Pi(s)))$$

ψ satisfies property P5

To prove Property (P5) for the given function $\psi(\mu(\Pi(s)))$, we will follow these steps:

Prove (i): Show that if $\mu(\Pi(E)) = \mu(\Pi(E'))$, then $\psi(\mu(\Pi(E))) = \psi(\mu(\Pi(E')))$ for all events $E, E' \in \mathcal{P}^i$. Prove (ii): Show that for every event $E \in \mathcal{P}^i \bigcap \mathcal{P}^j$, there exists a unique $y \in \mathbb{R}$ such that $\psi(\mu(\Pi(E))) = y$. Proof:

Prove (i): Given that $\mu(\Pi(E)) = \mu(\Pi(E'))$, we want to show that $\psi(\mu(\Pi(E))) = \psi(\mu(\Pi(E')))$.

Let's evaluate the function $\psi(\mu(\Pi(E)))$ and $\psi(\mu(\Pi(E')))$:

$$\psi(\mu(\Pi(E))) = \mathbb{E}_\mu(\Pi(E))^{1-\gamma\cdot\sigma_\mu(\Pi(E))} \left(\frac{1+\eta}{|\mathcal{P}^i|}\right)^{\gamma\cdot\sigma_\mu(\Pi(E))}$$

$$\psi(\mu(\Pi(E'))) = \mathbb{E}_\mu(\Pi(E'))^{1-\gamma\cdot\sigma_\mu(\Pi(E'))} \left(\frac{1+\eta}{|\mathcal{P}^i|}\right)^{\gamma\cdot\sigma_\mu(\Pi(E'))}$$

Since $\mu(\Pi(E)) = \mu(\Pi(E'))$, it follows that $\mathbb{E}_\mu(\Pi(E)) = \mathbb{E}_\mu(\Pi(E'))$ and $\sigma_\mu(\Pi(E)) = \sigma_\mu(\Pi(E'))$. Therefore, $\psi(\mu(\Pi(E))) = \psi(\mu(\Pi(E')))$, which proves part (i) of Property (P5).

Prove (ii): Let $E \in \mathcal{P}^i \bigcap \mathcal{P}^j$. We want to show that there exists a unique $y \in \mathbb{R}$ such that $\psi(\mu(\Pi(E))) = y$.

Since $\psi(\mu(\Pi(s)))$ is a continuous function of $\mathbb{E}_\mu(\Pi(s))$ and $\sigma_\mu(\Pi(s))$, and both $\mathbb{E}_\mu(\Pi(s))$ and $\sigma_\mu(\Pi(s))$ are continuous functions of $\mu(\Pi(s))$, it follows that $\psi(\mu(\Pi(s)))$ is a continuous function of $\mu(\Pi(s))$.

Now, consider the function $\psi(\mu(\Pi(E)))$:

$$\psi(\mu(\Pi(E))) = \mathbb{E}_\mu(\Pi(E))^{1-\gamma\cdot\sigma_\mu(\Pi(E))} \left(\frac{1+\eta}{|\mathcal{P}^i|}\right)^{\gamma\cdot\sigma_\mu(\Pi(E))}$$

Since $\psi(\mu(\Pi(s)))$ is a continuous function of $\mu(\Pi(s))$, and given that $E \in \mathcal{P}^i \bigcap \mathcal{P}^j$, it follows that there exists a unique $\mu(\Pi(E))$ for the event E. This, in turn, determines unique values for $\mathbb{E}_\mu(\Pi(E))$ and $\sigma_\mu(\Pi(E))$, as they are continuous functions of $\mu(\Pi(s))$.

Now, let's evaluate the function $\psi(\mu(\Pi(E)))$ again:

$$\psi(\mu(\Pi(E))) = \mathbb{E}_\mu(\Pi(E))^{1-\gamma\cdot\sigma_\mu(\Pi(E))} \left(\frac{1+\eta}{|\mathcal{P}^i|}\right)^{\gamma\cdot\sigma_\mu(\Pi(E))}$$

Since the values of $\mathbb{E}_\mu(\Pi(E))$ and $\sigma_\mu(\Pi(E))$ are unique for the given $\mu(\Pi(E))$, the value of $\psi(\mu(\Pi(E)))$ is also unique for the given event E. Let this unique value be $y \in \mathbb{R}$.

Thus, we have shown that for every event $E \in \mathcal{P}^i \bigcap \mathcal{P}^j$, there exists a unique $y \in \mathbb{R}$ such that $\psi(\mu(\Pi(E))) = y$. This proves part (ii) of Property (P5).

Proposition 2: ψ satisfies properties P1, P2, P3, P4, P5 and P6.

Proof:

In order to prove that a function $\psi : \mu \to \mathbb{R}_+$ that satisfies properties (P1), (P2), (P3), (P4), (P5), and (P6). For degenerate μ, P6 is inconsequential. Therefore we only have to prove the case of non-degenerate μ.

Apply properties (P5) and (P6) to establish the relationship between $\psi(\mu(\Pi(E)))$ and $\lambda\mathbb{E}\mu(\Pi(E)) + (1-\lambda)\frac{1}{|\mathcal{P}^i|}$. Use the definitions of the expectations and standard deviations for the marginal compound probabilities. Show that the desired result holds for any non-degenerate μ and event $E \in \mathcal{P}^i$. Proof:

Then, we follow the steps below

Step 1: Applying properties (P5) and (P6)

By Property (P5), for all events $E, E' \in \mathcal{P}^i$, if $\mu(\Pi(E)) = \mu(\Pi(E'))$, then $\psi(\mu(\Pi(E))) = \psi(\mu(\Pi(E')))$.

By Property (P6), if $|\mathcal{P}^i| < |\mathcal{P}^j|$ and $\Psi(\mu_1(\Pi(E_i))) \geq \Psi(\mu_2(\Pi(E_i)))$, then $\Psi(\mu_1(\Pi(E_j))) \geq \Psi(\mu_2(\Pi(E_j)))$.

Step 2: Using the definitions of expectations and standard deviations for the marginal compound probabilities

Given the definitions of $\mathbb{E}_\mu(\Pi(s))$ and $\sigma(\mu(\Pi(s)))$, we can rewrite them as:

$\mathbb{E}\mu(\Pi(s)) = \sum^N i = 1p(\pi_i) \times \pi_i(s)$

$\sigma(\mu(\Pi(s))) = \sqrt{\sum_{i=1}^N p(\pi_i)(\pi_i(s) - \mathbb{E}_\mu(\Pi(s)))}$

Step 3: Showing that the desired result holds for any non-degenerate μ and event $E \in \mathcal{P}^i$

Now, let's assume that ψ satisfies properties (P1), (P2), (P3), (P4), (P5), and (P6). We want to show that for any non-degenerate μ and event $E \in \mathcal{P}^i$, there exists $\lambda \in [0, 1]$ such that $\Psi(\mu(\Pi(E))) = \lambda\mathbb{E}\mu(\Pi(E)) + (1 - \lambda)\frac{1}{|\mathcal{P}^i|}$.

By applying properties (P5) and (P6), we know that there is a unique $y \in \mathbb{R}$ for which $\psi(\mu(\Pi(E))) = y$. Moreover, if $|\mathcal{P}^i| < |\mathcal{P}^j|$, then $\Psi(\mu_1(\Pi(E_i))) \geq \Psi(\mu_2(\Pi(E_i)))$ implies $\Psi(\mu_1(\Pi(E_j))) \geq \Psi(\mu_2(\Pi(E_j)))$.

Let $\lambda = \frac{\Psi(\mu(\Pi(E))) - \frac{1}{|\mathcal{P}^i|}}{\mathbb{E}\mu(\Pi(E)) - \frac{1}{|\mathcal{P}^i|}}$.

We can rearrange this equation to obtain the desired result:

$\Psi(\mu(\Pi(E))) = \lambda\mathbb{E}\mu(\Pi(E)) + (1 - \lambda)\frac{1}{|\mathcal{P}^i|}$

Now, we need to show that $\lambda \in [0, 1]$.

Since Ψ is a function mapping to $\mathbb{R}_+$, we have $\Psi(\mu(\Pi(E))) \geq 0$.

Appendix B: An Example of Ambiguity Adjustment Function

A parametric example of an Ambiguity Adjustment Function, that satisfies properties (P1) through (P5) is given by:

$$\psi'(\mu(\Pi(E))) = \mathbb{E}_\mu(\Pi(E))^{1-\alpha\cdot\sigma_\mu(\Pi(E))}$$

where $\alpha \in (-2, 2)$. Here, α represents the DM's ambiguity attitude, so that $\alpha < 0$ indicates ambiguity seeking behavior and $\alpha > 0$ indicates ambiguity aversion. $\alpha = 0$ is the case for an ambiguity neutral DM. This function is very similar to the one case proposed in the Gajdos et al. (2008) model, but without the "pessimistic criterion" that the DM overweights not only the states where information is more precise, but also the worst case scenarios in terms of utility. Here we only require that the DM alters the weight of events that have more imprecise information about their probability, and that can happen either for "good" or "bad" outcomes. This is also in the same spirit as Klibanoff et al. (2005) Φ function, in that the function is convex(concave) whenever the DM is ambiguity seeking(averse), no matter the outcome. But it also retains the same lack of meaningful predictions about what conditions what determines ambiguity attitude: we can estimate values for α in different situations, but this initial version of the model does not imply any meaningful predictions about how ambiguity attitudes may change due to changes in the probability distributions of events, their outcome domains, and other characteristics of the choice set.

That is the main reason for us to adopt a framework of elaborating postulates regarding how decision-makers deal with event probabilities (and not directly stating anything about the preferences of acts contingent on ambiguous events themselves), and also why we introduce postulate P6, resulting in a model that produces meaningful predictions about ambiguity attitudes that match the empirical literature.

Appendix C: A Note on Continuous State-Spaces

In this paper we purposely focus on discrete state-space problem. When dealing with decision under ambiguity, these seem to be the most relevant cases, since the simplification of the possible events and their associated outcomes is a common way for individuals to deal with complex and/or incomplete information about probabilities of events. Even in the experimental literature, it is usual to represent continuous distributions, even some of the most known ones as the univariate normal distribution as discrete approximations (Lian et al., 2019), since the continuous distribution information itself may be too complex for the decision-maker to form meaningful scenarios that she can use in choosing from a given set of acts. There has also been a long-standing literature on simplifying rules used by decision-makers, when faced with complex decisions (Kahneman et al., 1982; Sundstroem, 1987), without losing significant effectiveness in decision-making (Bruce & Johnson, 1996; Hertwig & Todd, 2003)

However, we also recognize that there may be instances where considering a continuous state-space may be useful, specially to relate discrete and continuous spaces through measure theory. Therefore, we alter postulates P5 and P6 to adapt the definitions of ambiguity adjustment function and the context-based ambiguity adjustment function, but now for continuous finite state-spaces. Basically, now the use of the cardinality of an Act-Induced Partition now does not make any sense (since cardinality is not a good measure of how likely an event in a continuous state-space is). Instead, we use the Lebesgue Measure to measure it. Below we adapt postulates P5 and P6 considering that, which we rename postulates P5' and P6', respectively.

P5'(Belief Symmetry - Continuous State-Spaces) Let $\mathcal{P}^i = \bigcup_{i=1}^n E_i$ and $\mathcal{P}^j = \bigcup_{j=1}^n E_j$ be act-induced partitions for some acts f_i and f_j, respectively. Then, the following holds:

(i) For all events $E, E' \in \mathcal{P}^i$, $\mu(E) = \mu(E') \implies \Psi(\mu(E)) = \Psi(\mu(E'))$.

(ii) For every event $E \in \mathcal{P}^i \bigcap \mathcal{P}^j$ and $l(\mathcal{P}^i) = l(\mathcal{P}^j)$ (where $l(A)$ denotes the Lebesgue measure of a set A), there is a unique $y \in \mathbb{R} : \Psi(\mu(E)) = y$.

P6'(Partition Monotonicity - Continuous State-Spaces) For any non-degenerate probability distribution μ_1 and μ_2 such that $\mu_1(\Pi(E_i)) = \mu_2(\Pi(E_i))$ for some $E_i \in \mathcal{P}^i$ and $\mu_1(\Pi(E_j)) = \mu_2(\Pi(E_j))$ for some $E_j \in \mathcal{P}^j$, if $l(\mathcal{P}^i) < l(\mathcal{P}^j)$, then $\Psi(\mu_1(\Pi(E_i))) \geq \Psi(\mu_2(\Pi(E_i)))$ implies $\Psi(\mu_1(\Pi(E_j))) \geq \Psi(\mu_2(\Pi(E_j)))$.

Analogously, we also redefine our parametric example function ψ', but now adapted to a continuous state-space:

$$\psi(\mu(\Pi(E))) = \mathbb{E}_\mu[E]^{(1-\gamma\cdot\sigma_\mu(E_s^i))} \left(\frac{l(E)+\eta}{\sum_i l(E_i)}\right)^{\gamma\cdot\sigma_\mu(E_s^i)} \tag{5}$$

where $l(E)$ is the Lebesgue measure of an event $E \in \mathcal{P}^i$ for an act f_i being evaluated, and $\sum_i l(E_i)$ is the sum of lebesgue measures for all events $E_i \in \mathcal{P}^i$, including the event E itself. The proof that this function is a Context Ambiguity Adjustment Function (for continuous state-spaces) is analgous to the discrete case proof of Appendix A.

References

ABDELLAOUI, M.; BAILLON, A.; PLACIDO, L.; WAKKER, P. P. The rich domain of uncertainty: Source functions and their experimental implementation. American Economic Review, v. 101, n. 2, p. 695-723. 2011.

AHN, D. S. Ambiguity without a state space. The Review of Economic Studies, v. 75, n. 1, p. 3-28, 2008.

ALMEIDA, S.; RANGEL, M. Tıtulo: Probabilistic Sophistication, Sources of Uncertainty, and Cognitive Ability: Experimental Evidence. 2014.

ARDALAN, K. Equity home bias: A review essay. Journal of Economic Surveys, v. 33, n. 3, p. 949-967, 2019.

ATTANASI, G.; GOLLIER, C. MONTESANO, A.; PACE, N. Eliciting ambiguity aversion in unknown and in compound lotteries: a smooth ambiguity model experimental study. Theory and decision, v. 77, p. 485-530, 2014.

BORDALO, P.; GENNAIOLI, N.; SHLEIFER, A. Salience theory of choice under risk. The Quarterly journal of economics, n. 127, v.3, p. 1243-1285. 2012.

BORDALO, P.; GENNAIOLI, N.; SHLEIFER, A. Salience. National Bureau of Economic Research. 2021.

BRUCE, Alistair C.; JOHNSON, Johnnie EV. Decision-making under risk: Effect of complexity on performance. Psychological Reports, v. 79, n. 1, p. 67-76, 1996.

CAMERER, C.; WEBER, M. Recent developments in modeling preferences: Uncertainty and ambiguity. Journal of risk and uncertainty, v.5, n.4, p. 325-370. 1992.

CHATEAUNEUF, A.; EICHBERGER, J.; GRANT, S.. Choice under uncertainty with the best and worst in mind: Neo-additive capacities. Journal of Economic Theory, v. 137, n. 1, p. 538-567, 2007.

CHATEAUNEUF, A.; FARO, J. H. Ambiguity through confidence functions. Journal of Mathematical Economics, v. 45, n. 9-10, p. 535-558, 2009.

CHEW, S. H.; EBSTEIN, R. P.; ZHONG, S. Ambiguity aversion and familiarity bias: Evidence from behavioral and gene association studies. Journal of Risk and Uncertainty, 44(1), 1-18. 2012.

CHEW, S. H.; LI, K. K.; CHARK, R.; ZHONG, S. Source preference and ambiguity aversion: Models and evidence from behavioral and neuroimaging experiments. In Neuroeconomics. Emerald Group Publishing Limited. 2008

CHEW, S. H.; LI, K. K.; SAGI, J. S. Experimental Evidence of Source Preference: Familiarity and Home Bias. Available at SSRN 3870716. 2021.

CHEW, S. H.; MIAO, B.; ZHONG, S. Partial ambiguity. Econometrica, v. 85, n.4, p. 1239-1260. 2017

CONTE, A.; HEY, John D. Assessing multiple prior models of behaviour under ambiguity. Journal of Risk and Uncertainty, v. 46, p. 113-132, 2013.

DE LARA RESENTE, J. G.; WU, G. Competence effects for choices involving gains and losses. Journal of Risk and Uncertainty, v. 40, n. 2, p. 109-132. 2010

DEQUECH, D. Fundamental uncertainty and ambiguity. Eastern economic journal, n. 26, v.1, p. 41-60. 2000.

DERTWINTEL-KALT, M.; FRAY, J.; KOESTER, M.. Optimal Stopping in a Dynamic Salience Model. Mimeo, 2021.

DI MAURO, C.; MAFFIOLETTI, A. An experimental investigation of the impact of ambiguity on the valuationof self-insurance and self-protection. Journal of Risk and Uncertainty, 13(1), 53-71. 1996.

DIMMOCK, S. G.; KOUWENBERG, R.; MITCHELL, O. S., PEIJNENBURG, K. Estimating ambiguity preferences and perceptions in multiple prior models: Evidence from the field. Journal of Risk and Uncertainty, 51(3), 219-244. 2015.

DIMMOCK, S. G.; KOUWENBERG, R.; WAKKER, P. P. Ambiguity attitudes in a large representative sample. Management Science, n. 62, v.5, 1363-1380. 2016.

ANTONIOU, Constantinos; HARRIS, Richard DF; ZHANG, Ruogu. Ambiguity aversion and stock market participation: An empirical analysis. Journal of Banking Finance, v. 58, p. 57-70, 2015.

DIMMOCK, S. G.; KOUWENBERG, R.; MITCHELL, O.S.; PEIJNENBURG, K. Ambiguity aversion and household portfolio choice puzzles: Empirical evidence. Journal of Financial Economics, v. 119, n. 3, p. 559-577, 2016.

DURSCH, P.; ROMER, D.; ROTH, B. Intertemporal stability of ambiguity preferences. Discussion Paper Series (No. 548). 2013.

ELLSBERG, D. Risk, ambiguity, and the Savage axioms. The quarterly journal of economics, v. 75, n. 4 p. 643-669. 1961.

EPSTEIN, L. G. A definition of uncertainty aversion. The Review of Economic Studies, v. 66, n. 3, p. 579-608. 1999.

EPSTEIN, L. G.; ZHANG, J. Subjective probabilities on subjectively unambiguous events. Econometrica, n. 69, pp. 265-306. 2001.

FOX, Craig R.; TVERSKY, Amos. Ambiguity aversion and comparative ignorance. The quarterly journal of economics, v. 110, n. 3, p. 585-603, 1995.

GHIRARDATO, Paolo. Defining ambiguity and ambiguity attitude. Uncertainty in Economic Theory: Essays in Honor of David Schmeidler's 65th Birthday, v. 63, p. 36-45, 2004.

GHIRARDATO, Paolo; MACCHERONI, Fabio; MARINACCI, Massimo. Differentiating ambiguity and ambiguity attitude. Journal of Economic Theory, v. 118, n. 2, p. 133-173, 2004.

GILBOA, I.; SCHMEIDLER, D. Maxmin Expected Utility with a Non-unique Prior. Journal of Mathematical Economics 18. p. 141–153. 1989.

GNEEZY, U.; IMAS, A.; LIST, John. Estimating individual ambiguity aversion: A simple approach. National Bureau of Economic Research, 2015.

GUTIERREZ, C.; KEMEL, E. Measuring natural source dependence. Working Paper. 2021.

HEATH, C.; TVERSKY, A. Preference and belief: Ambiguity and competence in choice under uncertainty. Journal of risk and uncertainty, 4(1), 5-28. 1991.

HENSHER, D. A.; LI, Z.; HO, C. The role of source preference and subjective probability in valuing expected travel time savings. Travel Behaviour and Society, 2(1), 42-54. 2015.

HERTWIG, Ralph; TODD, Peter M. More is not always better: The benefits of cognitive limits. Thinking: Psychological perspectives on reasoning, judgment and decision making, p. 213-231, 2003.

KAHNEMAN, D.; SLOVIC, S. P.; SLOVIC, P.; TVERSKY, A. Judgment under uncertainty: Heuristics and biases. Cambridge university press, 1982.

KAHNEMAN, D.; TVERSKY, A. Prospect theory: An analysis of decision under risk. Econometrica, v. 47, v. 2, p. 263-292. 1979.

TVERSKY, A.; KAHNEMAN, D. Advances in prospect theory: Cumulative representation of uncertainty. Journal of Risk and uncertainty, v. 5, p. 297-323, 1992.

KARNI, E. Savages' Subjective Expected Utility Model. Retrieved March, v. 7, p. 2010-2021, 2005.

KLIBANOFF, P.; MARINACCI, M.; MUKERJI, S. A smooth model of decision making under ambiguity. Econometrica, n. 73, v.6, p. 1849-1892. 2005.

KLIBANOFF, P.; MARINACCI, M.; MUKERJI, S. Definitions of ambiguous events and the smooth ambiguity model. Economic Theory, n. 48, v.2, p. 399-424. 2011.

MACHINA, M. J.; SCHMEIDLER, D. A more robust definition of subjective probability. Econometrica: Journal of the Econometric Society, v. 60, n. 4, p. 745-780. 1992.

MACHINA, M. J. Risk, Ambiguity and the Rank-Dependence Axioms. American Economic Review, v. 99, n. 1, p. 385-392, 2009.

NIELSEN, C.; SEBALD, A.; SORENSEN, P. Testing for Salience Effects in Choice under Risk. Mimeo. 2021.

RIEGER, M. O.; WANG, M.. Can ambiguity aversion solve the equity premium puzzle? Survey evidence from international data. Finance Research Letters, v. 9, n. 2, p. 63-72, 2012.

SAVAGE. The Foundations of Statistics. Wiley, New York 1954.

SCHMEIDLER, D. Subjective probability and expected utility theory without additivity. Econometrica v. 57, p. 571–587. 1989.

STARMER, C. Developments in non-expected utility theory: The hunt for a descriptive theory of choice under risk. Journal of economic literature, v. 38, n. 2, p. 332-382. 2000.

SUNDSTRÖM, Gunilla A. Information search and decision making: The effects of information displays. Acta Psychologica, v. 65, n. 2, p. 165-179, 1987.

TRAUTMANN, S. T.; VAN DE KUILEN, G. Ambiguity attitudes. The Wiley Blackwell handbook of judgment and decision making, v.1, 89-116. 2015.

TRAUTMANN, S.; WAKKER, P. Making the Anscombe-Aumann approach to ambiguity suited for descriptive applications. Working Paper, Erasmus University. 2012.

WAKKER, P. P. Prospect theory: For risk and ambiguity. Cambridge university press. 2010.

ZHANG, J. Subjective ambiguity, expected utility and choquet expected utility. Economic Theory, n. 20, pp. 159-181. 2002.