Counterfactual Priors: A Bayesian Response to Ellsberg's Paradox

Koundouri, Phoebe and Pittis, Nikitas and Samartzis, Panagiotis

February 2023

Counterfactual Priors: A Bayesian Response to Ellsberg's Paradox

Phoebe Koundouri*† Nikitas Pittis‡ Panagiotis Samartzis*§

Abstract

This paper analyzes the root cause of Ellsberg-type choices. This class of problems share the feature that at the time of the decision, $t = m$, the decision maker (DM) possesses partial information, $\mathbf{I}_m$, about the events/propositions of interest $\mathcal{F}$: DM knows the objective probabilities of the sub-class $\mathcal{F}_1$, $\mathcal{F}_1 \subset \mathcal{F}$ only, whereas she is uninformed about the probabilities of the complement $\mathcal{F}_1'$. As a result, DM may slip into the state of "comparative ignorance" (see Heath and Tversky 1991 and Fox and Tversky 1995). Under this state, DM is likely to exhibit "ambiguity aversion" (AA) for the events of $\mathcal{F}_1'$ relative to those of $\mathcal{F}_1$. AA, in turn results in DM having non-coherent beliefs, that is, her prior probability function, $P_0^{\mathbf{I}_m}$, is not additive. A possible way to mitigate AA is to motivate DM to form her prior in a state of "uniform ignorance". This may be accomplished by inviting DM to bring herself to the hypothetical time $t = 0$, in the context of which $\mathbf{I}_m$ was still a contingency, and trace her "counterfactual prior", P_0^c, "back then". Under uniform ignorance, DM may adhere to the "Principle of Indifference", thus identifying P_0^c with the uniform distribution. Once P_0^c is elicited, DM can embody the existing information $\mathbf{I}_m$ into her current, actual set of beliefs P_m by means of Bayesian Conditionalization. In this case, we show that P_m is additive.

Keywords: counterfactual priors, ambiguity, ellsberg paradox.

JEL Classification: C44, D81, D83, D89

---

*These authors acknowledge funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 951424; ERC-SyG Water-Futures).

†School of Economics and ReSEES Laboratory, Athens University of Economics and Business; Sustainable Development Unit, Athena Research Center; World Academy of Arts and Science; UN SDSN-Europe; e-mail: pkoundouri@aueb.gr, phoebe.koundouri@icre8.eu (corresponding author)

‡Department of Banking and Financial Management, University of Piraeus, Greece; International Center for Research on the Environment and the Economy, Greece; e-mail: npittis@unipi.gr

§Department of Economics, University of Macedonia, Greece; School of Economics, Athens University of Economics and Business, Greece; e-mail: psamartzis@uom.edu.gr

1 Introduction

Consider a decision maker, DM, who is about to make a decision (for the first time in her epistemic life) at period $t = m$ (now), that is to choose an act, f, from a set of available acts, $\mathcal{A}$. In order to make this decision she must assign probabilities to the events/propositions, A, of the space $\mathcal{F}$ that affect the outcomes of her actions. Throughout the paper we assume that the DM possesses partial information, $\mathbf{I}_m$, about the events of $\mathcal{F}$. Specifically, the DM knows the objective probabilities of the events in the sub-class $\mathcal{F}_1$, $\mathcal{F}_1 \subset \mathcal{F}$ only. On the contrary, she is completely uninformed about the probabilities of the events in $\mathcal{F}_1' = \mathcal{F} \smallsetminus \mathcal{F}_1$. This information asymmetry is characteristic of the epistemic situation, hereafter referred to as ES, underlying the so-called "Ellsberg choices" (Ellsberg 1961). Specifically, in the context of the well-known "three-colors-one-urn" case (see next section), $\mathbf{I}_m$ is the proposition that "there are 30 red balls inside the urn". More specifically, the DM is faced with an urn containing 90 balls, 30 red, and an unknown number of black and yellow balls. Under $\mathbf{I}_m$, the DM prefers betting on red to betting on black and betting on black or yellow to betting on red or yellow, thus contradicting the Savage's Sure-Thing Principle. A direct implications of these choices is that the DM's subjective probability function is not additive (or coherent).

The main question addressed in this paper is the following: Given that the DM is dealing with the epistemic situation ES at $t = m$, how should she form her prior subjective probability function P_0 defined on $\mathcal{F}$? An important feature of ES is that there is no actual information-free time point, say $t = 0$, in the DM's epistemic life: The time at which the DM becomes interested in the phenomenon for the first time is $t = m$ at which she already possesses information $\mathbf{I}_m$. This means that the DM does not have the option of forming an actual information-free prior P_0 (an $ur - prior$). Hence, the DM must decide how to handle $\mathbf{I}_m$ in the process of forming her prior subjective probability function. To that end, the DM has the following two options:

(i) The first option is to build her prior probability function, $P_0^{\mathbf{I}_m}$, under the direct influence of $\mathbf{I}_m$ at $t = m$. In such a case, $P_0^{\mathbf{I}_m}(\mathbf{I}_m) = 1$.

(ii) The second option involves a counterfactual move: The DM is invited to "mentally travel back in time", at the hypothetical time $t = 0$, which corresponds to a "tabula rasa" epistemic state, in the context of which $\mathbf{I}_m$ was still a contingency, and trace her counterfactual prior, P_0^c, "back then". This means that P_0^c is the prior that the DM would have had, if $\mathbf{I}_m$ was not known (to her). In this case, $P_0^c(\mathbf{I}_m) = p < 1$. In this counterfactual epistemic state the proposition $\mathbf{I}_m$ is treated as one of many unrealized possibilities, that is, it is on a par with every other conceivable proposition $\mathbf{I}_m'$. Once P_0^c is elicited, the DM can embody the existing information $\mathbf{I}_m$ into her current, actual set of beliefs P_m (those at $t = m$ on which her decision is based), by means of Bayesian Conditionalization (BC) using P_0^c as the appropriate vehicle of conditionalization, that is $P_m(A) = P_0^c(A \mid \mathbf{I}_m)$, $A \in \mathcal{F}$.

The DM's choice between $P_0^{\mathbf{I}_m}$ and P_0^c is likely to have important implications for the "additivity properties" of her prior probability function. More specifically, we raise the following question: which of the aforementioned two probability functions is more likely to obey the rules of formal probability calculus? Put differently, which of the two ways of forming a prior, the actual or the counterfactual, is more conducive to Bayesian rationality? Take $P_0^{\mathbf{I}_m}$ first. In the context of $P_0^{\mathbf{I}_m}$, the DM builds her prior probabilities under the influence of the partial information $\mathbf{I}_m$ which is tantamount to saying that the DM operates in the epistemic context of "comparative ignorance" (see Heath and Tversky 1991 and Fox and Tversky 1995). Specifically, the DM feels more ignorant about the objective probabilities of $\mathcal{F}_1'$ than those of $\mathcal{F}_1$. As a result, the DM is likely to exhibit "ambiguity aversion" for the events of $\mathcal{F}_1'$ relative to those of $\mathcal{F}_1$. Ambiguity aversion, in turn results in the DM having non-coherent beliefs, that is $P_0^{\mathbf{I}_m}$ is not additive.

On the other hand, in the context of P_0^c, the DM develops her probabilistic beliefs in the context of "uniform ignorance", that is she is as agnostic about the probabilities of $\mathcal{F}_1$ as she is about the probabilities of $\mathcal{F}_1'$. In such a "counterfactually symmetric" epistemic framework, the DM might not be susceptible to Ellsberg-type relative ambiguity, since her probabilistic knowledge on $\mathcal{F}_1'$ is no longer inferior to that on $\mathcal{F}_1$. For example, Fox and Tversky (1995) provide empirical evidence suggesting that the DM's ambiguity aversion decreases or even disappears in a non-comparative environment of uniform ignorance (see also Chow and Sarin 2001 for somewhat less supportive results for ambiguity vanishing). As a result, the DM may regain her Bayesian attitude, for example by identifying P_0^c with the uniform prior. Indeed, given that at this counterfactual epistemic state she does not possess any information ($\mathbf{I}_m$ is not assumed to be known), the adoption of the uniform prior is not arbitrary, but instead it represents the only appropriate way to describe DM's beliefs in the epistemic state of zero information. This is the view of the so-called Objective Bayesians, who argue that if the DM's knowledge about the possible (mutually exclusive and jointly exhaustive) outcomes is symmetric, then this symmetry should be reflected on the DM's assignment of probabilities to these outcomes. This "symmetry thesis", in turn, entails the Principle of Indifference (POI), which states that if the DM does not have any epistemic reason to differentiate her probabilistic assignments over the outcomes then she is compelled to assign these outcomes the same probability.[1] As Norton (2006) remarks: "...beliefs must be grounded in reasons, so that when there are no differences in reasons there should be no differences in beliefs" (2006, pp. 3-4). In the Ellsberg case, if the DM makes the counterfactual move, then she will find herself in an epistemic state in which she has no reason to assign the three colors, R(ed), B(lack) and Y(ellow),different probabilities; hence $P_0^c(R) = P_0^c(B) = P_0^c(Y) = 1/3$. Any prior probability distribution other than that implies information that the DM (counterfactually) does not have. On this view, POI serves as an objective guiding principle that instructs the DM how to form a unique prior. This means

[1] Another principle that supports the adoption of the uniform prior in the epistemic state of no-information is Jaynes's Maximum Entropy Principle. More specifically, the uniform prior is the (unique) probability function that maximizes *Shannon's Entropy* (see Shannon 1948, Jaynes 1957).

that the DM's lack of information about the objective probabilities of $\mathcal{F}$ does not necessarily brings the DM to the epistemic state of ambiguity with respect to her prior; POI offers the DM an intuitively appealing and logically sound way to avoid probabilistic paralysis. Finally, if the DM selects the uniform prior as P_0^c, then her current probability function P_m, that arises from P_0^c via conditionalization on $\mathbf{I}_m$, is also additive.

What are the empirical implications of the foregoing discussion for Ellsberg's paradox? As already mentioned, if the DM does not realize that she has two options for building her prior, namely the $\mathbf{I}_m$-driven $P_0^{\mathbf{I}_m}$ and the counterfactual P_0^c, (instead of $P_0^{\mathbf{I}_m}$ only), then she is likely to slip into ambiguity aversion and Ellsberg-type choices. On the other hand, if the DM is elucidated that the Bayesian option P_0^c is also available to her, then she might decide to choose P_0^c over $P_0^{\mathbf{I}_m}$. Which of the two choices decision makers tend to make (when both choices are explained to them) is an interesting subject for empirical investigation. To that end, if the DM insists on $P_0^{\mathbf{I}_m}$, once the alternative P_0^c has been adequately explained to her, then the Ellsberg-type behavior persists and the associated paradox proves to be a robust empirical regularity. In such a case, we are faced with two alternative interpretations of the paradox: Either to condemn the DM as "stubbornly irrational" or to admit that ambiguity aversion is a rational "trait" of the DM's behaviour (rather than a temporary irrational "state" generated by contingent circumstances) that needs to be accounted for. The second interpretation is adopted by the large literature on "ambiguity aversion" that purports to rationally explain Ellsberg-type choices/beliefs by means of axiomatic systems of preferences that relax some of Savage's axioms (especially, the Sure Thing Principle, see, for example Schmeidler 1989, Gilboa and Schmeidler 1989, Maccheroni et al. 2006 and Cerreia-Vioglio et al. 2011). If, however, the DM switches from Ellsberg-type choices/beliefs to Bayesian ones, once she becomes fully aware of the counterfactual option, then the Ellsberg paradox is eliminated from the empirical domain. In such a case, the paradox is dissolved rather than solved. In other words, the cognitive state that produces ambiguity aversion and Ellsberg-type choices might prove to be temporary, arising from the DM's failure to realize all of her Bayesian options. Nevertheless, despite its normative virtues (see Section 4), the question of whether the DM accepts the counterfactual strategy after it is presented to her remains an empirical matter. The rest of the paper is organized as follows: Section 2 discusses the applicability of POI in the context of Ellsberg's "three-colours-one-urn" decision problem, under the assumption that the DM has decided to switch to the counterfactual mode. It also draws a sharp distinction between the epistemic state of "uncertainty" and that of "ambiguity". Section 3 shows formally the main thesis of the paper, namely that if the DM adheres to the proposed counterfactual strategy of forming her prior, the Ellsberg-type behavior does not arise. Section 4 analyzes from a normative point of view, the merits of P_0^c relative to those of $P_0^{\mathbf{I}_m}$. In particular, it attempts to answer the question "what methodological and/or psychological reasons can be found for preferring P_0^c to $P_0^{\mathbf{I}_m}$?" Section 5 concludes the paper.

2 Principle of Indifference, Small versus Large Worlds, Ambiguity versus Uncertainty

If the DM finds the counterfactual strategy, described in the previous section appealing, then she is just one step away from forming a coherent prior. All that she has to do is to subscribe to POI, thus identifying P_0^c with the uniform distribution. However, it is well known that POI is far from being a universally accepted principle. The most serious argument against POI is that it leads to inconsistent probability distributions, with each of them being dependent on how the relevant sample space is partitioned. This problem was first identified by Bertrand (1889) and for this reason it is usually referred to as "Bertrand's paradox". However, this paradox is relevant only for the cases in which the sample space is uncountably infinite, as in the often-cited case of the "cube factory", put forward by van Fraassen (1989). Obviously, the Bertrand-type inconsistencies are not relevant for the Ellsberg case, in which the sample space $\Omega = \{R, B, Y\}$ is finite.[2]

Another objection against POI is the following: Assume that the DM is interested in assigning probabilities to the outcomes (H or T) of a coin toss. Consider the following cases: Case I: DM knows nothing about the coin, which means that she has no reasons to believe that H is more or less probable than T. Hence, by POI, she sets $P(H) = P(T) = 0.5$.[3] Case II: The DM is assumed to know that the coin is biased (although she does not know which of the two outcomes it favors). This means that she has reasons to exclude the value $P(H) = 0.5$ from the set of possible values that she may assign to H. In this case, (so the argument goes) POI is silent as to the values that DM should assign to H and T. At first glance, this argument seems to be convincing. However, a more careful analysis of the argument suggests the following: Let us think of $P(H)$ as a random variable, X, that takes values in the interval $[0, 1]$. The probability that $X = x$ is almost surely equal to zero, for every $x \in [0, 1]$, including the "special" value $x = 0.5$. Hence, the information that DM entertains in Case II is that a specific value in $[0, 1]$ (namely $x = 0.5$) is impossible. Compare this information with that of Case I. In case I, DM still knows that the probability that $X = 0.5$ is "almost surely" zero, that is she knows that the event $X = 0.5$ is "almost impossible". This means that DM's information gains as she moves from the epistemic state of Case I to that of Case II, is equal to the "difference" between the propositions "$X = 0.5$ is almost impossible" and "$X = 0.5$ is impossible". In other words, the information differential between Cases I and II is almost surely zero. This in turn implies that if the uniform distribution is approprate in the context of Case I, then it remains almost surely so in the

[2] The emergence of the Bertrand paradox even in the cases in which Ω is uncountably infinite is controversial. Jaynes (1968) introduces his "invariance condition" in an attempt to resolve Bertrand-type problems. Rosenkrantz (1982) argues that this attempt is successful.

[3] This case is radically different from that in which DM assigns equal probabilities to H and T on the basis of the information that "coins are usually fair". In this case, DM possesses empirical information (instead of being in an information-free epistemic state), thus making an inductive inference.

context of Case II.

A more formal argument takes the form of the following proposition: The uniform distribution is the one that maximizes entropy in the epistemic context of Case II. This proposition may be proved as follows: In general, entropy is a continuous function of the n-tuples $(p_1, p_2, ..., p_n)$, $H = -\sum_{i=1}^{n} p_i \ln(p_i)$. In our case, $n = 2, p_1 = P(H)$ and $p_2 = P(T)$. Given that the 2-tuples (p_1, p_2) lie in a compact subset of $\mathbb{R}^2$, there is a 2-tuple where entropy is maximized. We want to show that this occurs at $(1/2, 1/2)$ and nowhere else. Suppose, without loss of generality, that $P(H) < P(T)$. Then, there exists $\epsilon > 0$ such that $P(H)+\epsilon < P(T) - \epsilon$. We will show that the entropy of $(P(H) + \epsilon, P(T) - \epsilon)$ is larger than the entropy of $(P(H), P(T))$. Therefore, since entropy is maximized at some 2-tuple, it is uniquely maximized at the 2-tuple with $P(H) = P(T) = 1/2$. The entropy of $(P(H) + \epsilon, P(T) - \epsilon)$ minus the entropy of $(P(H), P(T))$ equals

$$(P(H) + \epsilon) \ln (P(H) + \epsilon) + (P(T) - \epsilon) \ln (P(T) - \epsilon) - P(H) \ln (P(H)) - P(T) \ln (P(T)) =$$

$$= \quad P(H) \ln \left(1 + \frac{\epsilon}{P(H)}\right) + P(T) \ln \left(1 + \frac{\epsilon}{P(T)}\right) + \epsilon \ln \left(\frac{P(T) - \epsilon}{P(H) + \epsilon}\right) > 0.$$

Therefore, the entropy is maximized when $P(H) = P(T) = 1/2$.

Of course, a much simpler response to the criticism against POI in the context of Case II is that the Ellsberg case does not belong to the class of epistemic cases encoded by Case II. More generally, Ellsberg's paradox is a typical example of a "small world" which is immune to the problems that may be relevant for "large worlds". Incidentally, the distinction between small and large worlds was first made by Savage himself, who found the idea of treating both worlds uniformly as "ridiculous" and "preposterous". It is not a sound methodological practice to take an argument that may have some force in a particular domain (namely that of uncountably infinite Ω or unequal probabilities) and apply it uncritically to a domain in which it is irrelevant.

It is important to note that Ellsberg's paradox was designed to unearth a new epistemic state, that of "ambiguity" (in which the DM knows the probabilities of $\mathcal{F}_1$ but not of $\mathcal{F}_1'$) . This state lies between the traditional states of "risk" (in which the DM knows the objective probabilities of all the elements of $\mathcal{F}$) and "uncertainty" (in which the DM does not know any of the objective probabilities of the elements of $\mathcal{F}$). On this view, "ambiguity" describes an epistemic state that is distinct from that described by "uncertainty", which means that the two terms should not be used interchangeably. On the contrary, Gilboa and Marinacchi (2016) do not adhere to such a distinction: "Today, the terms “ambiguity”, “uncertainty” (as opposed to “risk”), and “Knightian uncertainty” are used interchangeably to describe the case of unknown probabilities." (2016, footnote 8). But this is not what Ellsberg's "three-colors-one-urn" paradox was designed to capture. The problem of how the DM assigns probabilities to $\mathcal{F}$ under uncertainty was well-known long before 1961, year at which Ellsberg devised his paradox. The new situation that Ellsberg's paradox brought to light is the one in which the DM faces risk and uncertainty *within the same decision* *problem*, in the sense that she knows the probabilities of $\mathcal{F}_1$ but not those of $\mathcal{F}_1'$. In other words, the DM's simultaneous exposure to risk and uncertainty is the trigger that may cause her to display Ellsberg-type behavior.

Having said that, we must point out that part of the blame for the confusion surrounding the terms "uncertainty" and "ambiguity" must be put on Ellsberg himself. In his classic (1961) paper, before he introduces the "three-colors-one-urn" paradox, mentioned above, he discusses the following "two-colors-two-urns" case: The DM is faced with two urns, urn I and urn II. Urn I contains 100 red (R) and black (B) balls in a proportion unknown to the DM. For urn II, the DM is informed that it contains 50 red and 50 black balls. The DM contemplates the following "acts": f_R: "bet on R in urn I", f_B: "bet on B in urn I", g_R: "bet on R in urn II", g_B: "bet on B in urn II". Ellsberg invites the DM to think whether she prefers (i) f_R versus f_B, (ii) g_R versus g_B (iii) f_R versus g_R and (iv) f_B versus g_B. He argues that most decision makers are indifferent between f_R and f_B as well as between g_R and g_B. However, they tend to prefer g_R to f_R and g_B to f_B. Ellsberg argues that a DM who exhibits this set of choices has non-additive beliefs. Whether Ellsberg's argument is valid or not depends on the preferred reading of the "two-colors-two-urns" case. To that end, we may distinguish the following two alternative interpretations.

On the first interpretation, Ellsberg's argument is not valid. The reason is that the "two-colors-two-urns" case gives rise to two distinct decision problems, i.e. one in which the sample space (states of nature) is $\Omega_I = \{R_I, B_I\}$ and the other in which the relevant space is $\Omega_{II} = \{R_{II}, B_{II}\}$. This means that f_R and f_B are defined on Ω_I, whereas g_R and g_B are defined on Ω_{II}, which in turn implies that f_R and f_B are compared by means of the preference relation "$\succeq_I$" whereas g_R and g_B are compared by means of "$\succeq_{II}$". As a result, the preferences (i) and (ii) are well-defined whereas (iii) and (iv) are not. The problem with the "two-colors-two-urns" case is that two distinct decision problems are mixed into one. Indeed, the DM's "preferences" of g_R to f_R and g_B to f_B imply nothing more than that the DM prefers to know the probabilities of R and B than not. But this is hardly a paradox. Instead, it is a manifestation of a more general rational disposition of the DM to prefer "more information" to "less information".

In footnote 7 (page 651) Ellsberg remarks: "Note that in no case are you invited to choose both a color and an urn freely." This suggests a second interpretation, according to which the two decision problems, mentioned above are merged into one whose sample space is $\Omega = \{R_I, B_I, R_{II}, B_{II}\}$. By making this move, Ellsberg brings the epistemic states of risk and uncertainty under the same roof, and only then his aforementioned argument for the non-additivity of the DM's beliefs becomes valid. Of course, on this interpretation the "two-colors-two-urns" case becomes structurally similar to the "three-colors-one-urn" one, in the sense that the DM knows the objective probabilities of some but not all of the members of the Boolean algebra generated by Ω. The important point to note is that uncertainty *per se*, especially in cases in which POI is applicable, does not cause any "additivity problems" in the DM's system of beliefs. In order for such problems to emerge, elements of uncertainty and elements of risk must be jointly present in a single decision problem, and this is the case that we identify as "ambiguity" in this paper.

3 Counterfactual Prior and its Implications for Ellsberg-type Choices

Although Ellsberg's paradox is very well known, let us briefly recast it in our own notation. Consider an urn that contains 90 balls with three different colors. Suppose also, that the the DM being at time $t = m$ is given the specific information $\mathbf{I}_m$ which takes the form of the following proposition: "30 balls are red and the remaining 60 balls are either black or yellow in unknown proportion". The DM (who is about to draw a ball at random) is offered two pairs of choices/actions: (a) Choose between f and g, where

f: "a bet on red"

g: "a bet on black".

(b) Choose between f^* and g^*, where:

f^*: "a bet on red or yellow"

g^*: "a bet on black or yellow".

The following table contains the outcomes for each action and state of nature:

	red ball	black ball	yellow ball
f	100	0	0
g	0	100	0
f^*	100	0	100
g^*	0	100	100

Under the subjective expected utility maximization framework, the choice between actions f and g (as well as between f^* and g^*), is based on the calculation of the expected utility of the two actions. Since the prizes are exactly the same, it follows that the DM prefers f to g ($f \succ g$) if and only if she believes that drawing a red ball is more probable than drawing a black ball and vice versa. If the DM believes that drawing a red ball is more probable than drawing a black ball, then probabilistic coherence requires her to believe that drawing a red or yellow ball is more probable than drawing a black or yellow ball. Therefore, if the DM prefers f to g, then she prefers f^* to g^* (and vice versa).

When surveyed, however, most people strictly prefer f to g and g^* to f^*, thus violating the aforementioned prediction of the theory. Moreover, such a pair of choices imply that the DM's subjective probability function is not additive, which runs against the basic tenet of Bayesianism. What is a possible explanation of such behavior? As mentioned in Introduction, this behavior may be explained in terms of the "comparative ignorance" hypothesis: f is preferred to g because the objective probability of "red" as opposed to that of "black" is known. Similarly, g^* is preferred to f^* because the objective probability of "black or yellow" is also known. In other words, the DM's behavior is due to her preference to bet on events of $\mathcal{F}_1$ than on those of $\mathcal{F}_1'$. More specifically, the information $\mathbf{I}_m$ induces the following partition $(\mathcal{F}_1, \mathcal{F}_1')$ of the algebra of propositions $\mathcal{F}$:

$$\mathcal{F}_1 = \{s_R, s_{BY}, s_{RBY}, \bot\},$$
$$\mathcal{F}_1' = \{s_B, s_Y, s_{RB}, s_{RY}\},$$

where

$$\begin{aligned} s_R &: \text{"a red ball is drawn"} \\ s_B &: \text{"a black ball is drawn"} \\ s_Y &: \text{"a yellow ball is drawn"} \\ s_{RB} &: \text{"a red or a black ball is drawn"} \\ s_{RY} &: \text{"a red or a yellow ball is drawn"} \\ s_{BY} &: \text{"a black or a yellow ball is drawn"} \\ s_{RBY} &: \text{"a red or a yellow or a black ball is drawn"} \\ \bot &: \text{"no ball is drawn".} \end{aligned} \tag{1}$$

The probabilistic content of $\mathbf{I}_m$ takes the form of the following objective probabilities,

$$Ch(s_R) = \frac{1}{3}, Ch(s_{BY}) = \frac{2}{3}, Ch(s_{RBY}) = 1, Ch(\bot) = 0,$$

where $Ch(A)$ denotes the objective probability (chance) of proposition A.

Remark 1 *Using the Knightian distinction between "risk" and "uncertainty", the DM faces a risky situation (known probabilities supplied by $\mathbf{I}_m$) with respect to $\mathcal{F}_1$, whereas she operates in an environment of uncertainty (unknown probabilities) with respect to $\mathcal{F}_1'$. Should the DM distinguish between risk and uncertainty and especially should she prefer the former over the latter? According to strict Bayesians, the answer is negative: The DM is always able to ascertain her own subjective probabilities of $\mathcal{F}_1'$ which in combination with the known probabilities of $\mathcal{F}_1$ yield a proper subjective probability function over the whole of $\mathcal{F}$.[4] This means that ambiguity aversion sets in when the DM does not treat risk and uncertainty symmetrically; when she prefers the former epistemic state over the latter.*

Remark 2 *It is important to note that $\mathbf{I}_m$ refers to the objective probabilities of the elements of $\mathcal{F}_1$. Whether the DM endorses these objective probabilities as*

[4] In fact in some forms of radical subjectivism, the DM is allowed to ignore the furnished objective probabilities for $\mathcal{F}_1$, stick to her own probabilistic judgments for $\mathcal{F}_1$, and still be rational.

her own subjective probabilities is another question. Most philosophers agree that rationality dictates that the DM should conform to the "probability coordination principle", according to which the DM adopts as her own subjective probabilities the corresponding objective ones, provided that the latter are known (see, for example, Strevens 2017). In the analysis that follows we tacitly assume that the DM adheres to the aforementioned principle.

Let us now assume that the DM decides to implement the counterfactual strategy proposed in the paper. This means that the DM "mentally goes back" to the hypothetical time $t = 0$, in which $\mathbf{I}_m$ was not certain, but instead it was one of the many alternative pieces of information (information propositions) that the DM could receive at $t = m$. At that hypothetical moment, the DM deliberates her probabilistic assignments on $\mathcal{F}$ relativized only with respect to the background information, $\mathbf{I}_B$, available at that moment.[5] The important thing to note is that at that hypothetical moment, the DM is "uniformly ignorant" about the objective probabilities of $\mathcal{F}$. Hence, in contemplating P_0^c, the DM does not enter the cognitive state of comparative ignorance, which as already mentioned, is considered to be the main cause of ambiguity aversion.

Based on $\mathbf{I}_B$ alone, the DM knows that one of the following "theoretical propositions" is true:

$$\begin{aligned}
\mathcal{H}_{(0,0,90)} &: \text{"0 red, 0 black and 90 yellow balls"} \\
\mathcal{H}_{(0,1,89)} &: \text{"0 red, 1 black and 89 yellow balls"} \\
&. \\
&. \\
&. \\
\mathcal{H}_{(0,90,0)} &: \text{"0 red, 90 black and 0 yellow balls"} \\
&. \\
&. \\
&. \\
\mathcal{H}_{(90,0,0)} &: \text{"90 red, 0 black and 0 yellow balls"}.
\end{aligned}$$

Note that each of these theoretical propositions gives rise to a certain probability distribution of the three colors. Let $\mathbf{H}$ denote the set of all the aforementioned propositions. It is obvious that at the hypothetical moment $t = 0$, the DM does not know which proposition of $\mathbf{H}$ is the true one. As a result, she treats the

[5] What kind of information does I_B consist of? Let us answer this question by first clarifying what kind on information is not allowed to be part of I_B : Any kind of probabilistic information, namely either direct information on the probabilities of $\mathcal{F}$, such as "the number of red balls in the urn is 30", or indirect information of those probabilities, such as "in a long series of trials, the relative frequency of red draws is 30 percent". If such probabilistic information is excluded from I_B, then I_B is allowed to contain information about the broad features of the chance mechanism at hand. For example, part of I_B is the proposition that "there is an urn containing 90 balls", as well as the proposition that "the balls in the urn are red, yellow and black only" and also that "a ball will be drawn at random".

elements of $\mathbf{H}$ as part of the domain of P_0^c. This in turn implies that the DM's relevant algebra of propositions is not $\mathcal{F}$, but rather the extended space $\mathcal{F}_{ext}^0$, that includes, apart from the empirical propositions defined in (1) the theoretical propositions $\mathcal{H}_{\mathbf{i}}$, $\mathcal{H}_{\mathbf{i}} \in \mathbf{H}$, $\mathbf{i} \in \boldsymbol{I}$ (together with their conjunctions, disjunctions and negations) as well, where $\boldsymbol{I} = \{\mathbf{i} \in \mathbb{N}^3 : \mathbf{0} \leq \mathbf{i} \leq \mathbf{1} \times 90 \text{ and } \mathbf{1}' \times \mathbf{i} = 90\} \subset \mathbb{N}^3$, $\mathbf{i} = (i_R, i_B, i_Y)'$ which denotes the 3×1 vector that contains the numbers of red, black and yellow balls in the urn, respectively and $\mathbf{1} = (1,1,1)'$. The resulting propositional space $\mathcal{F}_{ext}$ is a Boolean algebra. It must be noted that although the DM is interested in the propositions of $\mathcal{F}_{ext}$, the space that contains the propositions of "betting interest" for the DM remains $\mathcal{F}$, $\mathcal{F} \subset \mathcal{F}_{ext}$.

At $t = 0$, the DM is equally uninformed about the elements of $\mathcal{F}_{ext}$. In this state it is quite natural to assume that the DM adopts the non-informative or uniform prior P_0^c, according to which each color has equal probability of being drawn.

Since P_0^c obeys the rules of probability calculus, it also satisfies the law of total probability, according to which $\forall a \in \{R, B, Y\}$,

$$P_0^c(s_a) = \sum_{\mathbf{i} \in \boldsymbol{I}} P_0^c(s_a \mid \mathcal{H}_{\mathbf{i}}) P_0^c(\mathcal{H}_{\mathbf{i}}). \tag{2}$$

To calculate $P_0^c(s_a \mid \mathcal{H}_{\mathbf{i}})$, i.e. the probability of drawing a $a-$colored ball conditional on the hypothesis $\mathcal{H}_{\mathbf{i}}$, it is convenient to define $\boldsymbol{I}_a^k \subset \boldsymbol{I}$, to be the subset of vectors for which the number of $a-$colored balls in the urn, is exactly k, where $a \in \{R, B, Y\}$, and $0 \leq k \leq 90$. Clearly, $P_0^c(s_a \mid \mathcal{H}_{\mathbf{i}})$ is non-zero if $\mathbf{i} \in \boldsymbol{I}_a^k$. Moreover, $P_0^c(s_a \mid \mathcal{H}_{\mathbf{i}}) = \frac{k}{90}$, if $\mathbf{i} \in \boldsymbol{I}_a^k$ and $card(\boldsymbol{I}_a^k) = 91 - k$, $\forall k = 0, ..., 90$ and $\forall a \in \{R, B, Y\}$. As a result,

$$\begin{aligned} P_0^c(s_a) &= \sum_{\mathbf{i} \in \boldsymbol{I}} P_0^c(s_a \mid \mathcal{H}_{\mathbf{i}}) P_0^c(\mathcal{H}_{\mathbf{i}}) = \\ &= \sum_{k=0}^{90} \sum_{\mathbf{i} \in \boldsymbol{I}_a^k} P_0^c(s_a \mid \mathcal{H}_{\mathbf{i}}) P_0^c(\mathcal{H}_{\mathbf{i}}) = \\ &= \sum_{k=0}^{90} \frac{k}{90} (91 - k) \frac{1}{4,186} = \frac{1}{3}, \forall a \in \{R, B, Y\}. \end{aligned}$$

Specifically,

$$P_0^c(s_R) = P_0^c(s_B) = \frac{1}{3} \text{ and } P_0^c(s_{RY}) = P_0^c(s_{BY}) = \frac{2}{3}.$$

At time $t = 0$, the DM assesses not only her unconditional subjective probabilities, $P_0^c(s_a)$, but the conditional ones $P_0^c(s_a \mid I_{\mathbf{s}})$ as well, for some information $I_{\mathbf{s}}$. As already mentioned, $I_{\mathbf{s}}$ might take the form of direct or indirect probabilistic information. Note that $I_{\mathbf{s}} \in \mathbf{H}$, which contains all the alternative "information scenarios" that may turn out to be the case. In our case, $I_{\mathbf{s}} = \mathbf{I}_m = \boldsymbol{I}_R^{30}$, with $\mathbf{I}_m$ being the proposition that "30 of the 90 balls are red".[6]

[6] Formally, $\mathbf{I}_m$ may be expressed as the disjunction of a subset of $H_{\mathbf{i}}$, namely $\mathbf{I}_m = \{H_{(30,0,60)} \vee H_{(30,1,59)} \vee ... \vee H_{(30,59,1)} \vee H_{(30,60,0)}\}$.

Another example of $I_{\mathbf{s}} = \boldsymbol{I}_B^{40}$ could be the information that "40 of the 90 balls are black". The important thing to note is that in order for the DM to complete the process of calculating P_0^c at $t = 0$, she has to judge all the conditional probabilities $P_0^c(s_a \mid I_{\mathbf{s}})$, $I_{\mathbf{s}} \in \mathbf{H}$ rather than only the specific conditional probability $P_0^c(s_a \mid \mathbf{I}_m)$. This is because in the context of her counterfactual reasoning, the factual proposition $\mathbf{I}_m$ must be treated on a par with any possible (but hypothetical) information scenario $I_{\mathbf{s}} \in \mathbf{H}$.

Let us now calculate the DM's conditional probabilities $P_0^c(s_a \mid \mathbf{I}_m)$.[7] Using (2) we have that $\forall a \in \{R, B, Y\}$:

$$P_0^c(s_a \mid \mathbf{I}_m) = \sum_{\mathbf{i} \in \boldsymbol{I}} P_0^c(s_a \mid \mathcal{H}_{\mathbf{i}} \wedge \mathbf{I}_m) P_0^c(\mathcal{H}_{\mathbf{i}} \mid \mathbf{I}_m),$$

where

$$P_0^c(s_a \mid \mathcal{H}_{\mathbf{i}} \wedge \mathbf{I}_m) = \begin{cases} P_0^c(s_a \mid \mathcal{H}_{\mathbf{i}}), & \mathbf{i} \in \boldsymbol{I}_R^{30} \\ 0, & \mathbf{i} \notin \boldsymbol{I}_R^{30} \end{cases}$$

and

$$P_0^c(\mathcal{H}_{\mathbf{i}} \mid \mathbf{I}_m) = \begin{cases} \frac{1}{61}, & \mathbf{i} \in \boldsymbol{I}_R^{30} \\ 0, & \mathbf{i} \notin \boldsymbol{I}_R^{30} \end{cases}.$$

Therefore,

$$\begin{aligned} P_0^c(s_a \mid \mathbf{I}_m) &= \sum_{\mathbf{i} \in \boldsymbol{I}} P_0^c(s_a \mid \mathcal{H}_{\mathbf{i}} \wedge \mathbf{I}_m) P_0^c(\mathcal{H}_{\mathbf{i}} \mid \mathbf{I}_m) = \\ &= \sum_{\mathbf{i} \in \boldsymbol{I}_R^{30}} P_0^c(s_a \mid \mathcal{H}_{\mathbf{i}}) \frac{1}{61}, \forall a \in \{R, B, Y\}. \end{aligned}$$

For $a = R$, the last equation becomes,

$$P_0^c(s_R \mid \mathbf{I}_m) = \sum_{\mathbf{i} \in \boldsymbol{I}_R^l} P_0^c(s_R \mid \mathcal{H}_{\mathbf{i}}) \frac{1}{61} = \frac{1}{3}.$$

Similarly, for the other two values of a we have, $P_0^c(s_B \mid \mathcal{H}_{\mathbf{i}}) = \frac{k}{90}$, if $\mathbf{i} \in \boldsymbol{I}_R^{30} \cap \boldsymbol{I}_B^k$ and $P_0^c(s_Y \mid \mathcal{H}_{\mathbf{i}}) = \frac{k}{90}$, if $\mathbf{i} \in \boldsymbol{I}_R^{30} \cap \boldsymbol{I}_Y^k$ and $card(\boldsymbol{I}_R^{30} \cap \boldsymbol{I}_B^k) = card(\boldsymbol{I}_R^{30} \cap \boldsymbol{I}_Y^k) = 1$, $\forall k = 0, ..., 60$. As a result,

$$\begin{aligned} P_0^c(s_a \mid \mathbf{I}_m) &= \sum_{k=0}^{60} \sum_{\mathbf{i} \in \boldsymbol{I}_R^{30} \cap \boldsymbol{I}_a^k} P_0^c(s_a \mid \mathcal{H}_{\mathbf{i}}) \frac{1}{61} = \\ &= \sum_{k=0}^{60} \frac{k}{90} \frac{1}{61} = \frac{1}{3}, a \in \{B, Y\}. \end{aligned}$$

[7] The procedure for calculating any other conditional probability $P_0^c(s_a \mid I_{\mathbf{s}})$, $I_{\mathbf{s}} \in \mathbf{H}$ is entirely similar.

It follows that,

$$P_0^c(s_R \mid \mathbf{I}_m) = P_0^c(s_B \mid \mathbf{I}_m) = P_0^c(s_Y \mid \mathbf{I}_m) = \frac{1}{3} \tag{3}$$

and

$$P_0^c(s_{RY} \mid \mathbf{I}_m) = P_0^c(s_{BY} \mid \mathbf{I}_m) = P_0^c(s_{RB} \mid \mathbf{I}_m) = \frac{2}{3}. \tag{4}$$

Once the DM repeats the procedure outlined above for all information scenarios $I_\mathbf{s} \in \mathbf{H}$, then the DM's formation of her own counterfactual prior P_0^c is completed.

Now it is time for the DM to exit the counterfactual mode of probabilistic thinking and mentally return to the actual time point $t = m$, to implement the second step of the counterfactual strategy, namely to adopt the counterfactual conditional probabilities, given in (3) and (4), as her current subjective probabilities for $t = m$. By doing so, the DM ends up with the following additive subjective probability function P_m, defined on $\mathcal{F}$:

$$\begin{aligned} P_m(s_R) &= P_m(s_B) = P_m(s_Y) = \frac{1}{3} \\ P_m(s_{RY}) &= P_m(s_{BY}) = P_m(s_{RB}) = \frac{2}{3}. \end{aligned}$$

It must be noted that the fact that new probability of s_R is equal to the corresponding old probability (equal to 1/3) is purely coincidental. If, for example, instead of $\mathbf{I}_m$ the actual information were $\mathbf{I}'_m$: "40 balls are red", then $P_0^c(s_R)$ would still be (under the uniform prior) equal to 1/3, but the new probability $P_m(s_R) = P_0^c(s_R \mid \mathbf{I}'_m)$ would now be equal to 4/9.

In fact, the results presented above can be easily generalized for any specific information $I_\mathbf{s}$ and any initial counterfactual prior, P_0^c (even if the DM does not adhere to the POI). This generalized result is stated in the form of the following proposition:

Proposition 3 *For any coherent (i.e. additive) counterfactual initial subjective probability function P_0^c, which assigns non-zero prior probabilities to each of the "theoretical propositions" in $\mathbf{H}$, and for any specific information $I_\mathbf{s}$, the subjective probability function P_m of time $t = m$, generated by $P_m(A) = P_0^c(A \mid I_\mathbf{s})$, $A \in \mathcal{F}$ is coherent (i.e. additive).*

Proof. See Appendix. ■

The above Proposition demonstrates the following: If the DM follows the two-step counterfactual way of processing any specific information that may come to know $t = m$, then her choices are of the Bayesian rather than of the Ellsberg type.

4 Normative Arguments for Counterfactual Priors

The following discussion pertains to the question of why should the DM choose P_0^c to $P_0^{\mathbf{I}_m}$, once the counterfactual option (P_0^c) has been articulated to her. More specifically, from the normative point of view, P_0^c fares better than $P_0^{\mathbf{I}_m}$ in (at least) two respects (for additional arguments in favor of P_0^c see Pittis et al. 2021):

(i) A rational DM must be able to locate the time point, $t = 0$, at which she builds her system of beliefs for the first time in the course of her epistemic life. Ideally, at this point no probabilistic information is available, in which case the DM forms her (actual) information-free prior P_0^a. The DM should also be able at $t = 0$ to elicit her prior conditional probabilities for any conceivable piece of information that may come in the future and commit herself that when in the future (i.e. at $t = m$) a specific proposition (say $\mathbf{I}_m$) proves to be true, she will adopt $P_0^a(A \mid \mathbf{I}_m)$ as her current probability $P_m(A)$, $A \in F$. In other words, a rational DM honors her ex-ante beliefs, thus being dynamically consistent. As already mentioned, in the context of ES, the actual time point $t = 0$ does not exist. So, if the DM finds it appealing to be able to track down the dynamic evolution of her beliefs, she must create this starting point counterfactually, thus forming P_0^c. On the other hand, in the context of the non-additive $P_0^{\mathbf{I}_m}$, there exists no singular, widely-accepted theory of how the DM should update her beliefs in the light of new information and what it means for the DM to be dynamically consistent.[8] Pahlke (2022), for example, summarizes the theoretical results of updating in the presence of ambiguity as follows: "Different updating rules are defined in the literature, but almost all such rules can lead to dynamically inconsistent behavior in combination with maxmin preferences." (2022, p. 86). In their critical survey of the ambiguity aversion literature Al-Najjar and Weinstein (2009) argue that the rationalization of Ellsberg's choices amounts to replacing one anomaly by other anomalies, namely one must accept as rational "decision makers who base their decisions on irrelevant sunk cost; update their beliefs based on taste, and not just information; have the ability to deform their beliefs at will; or express an aversion to information." (p. 250).

(ii) A rational DM should aim at each point in time to elicit her "true" beliefs rather than those driven by emotions or impulsive reactions. True beliefs are those that are robust to any further deliberation by the DM of the decision problem at hand. To borrow Al-Najjar and Weinstein's terminology, true beliefs are "immune to introspection" (2009, p. 252). This point was first raised by Rudolf Carnap (1962, 1971). More specifically Carnap argues that

[8] Gilboa and Scheidler 1993 suggest that a rational way for the DM to updated her non-Bayesian beliefs is according to the so-called "maximum likelihood update rule". This rule boils down to the Dempster–Shafer conditioning rule for preferences that can be simultaneously represented by a non-additive prior and by multiple priors (see Dempster 1968, Shafer 1976). Cheng (2022) introduces another updating rule for ambiguous beliefs represented by a set C of priors, the Relative Maximum Likelihood rule. This rule is based on applying Bayesian conditionalization to a properly defined subset of C.

the DM's pure or genuine probabilistic beliefs are expressed by what he calls "permanent dispositions to believe", which are identified with the DM's initial credence function P_0^c. This initial probability function stands in sharp contrast with the current probabilistic beliefs, $P_0^{\mathbf{I}_m}$, that the DM happens to have at some information-loaded point in time $t = m$. $P_0^{\mathbf{I}_m}$ does not capture the true probabilistic dispositions of DM, but instead they codify the DM's "momentary inclinations to believe" at time $t = m$. As already mentioned, in the Ellsberg case, these momentary beliefs are likely to be caused by the DM's "comparative ignorance". How do DM's current probabilistic beliefs, P_m inherit the "trait of the DM's underlying permanent intellectual character"? This can only be achieved if the DM conditionalizes on all the information accumulated between $t = 0$ and $t = m$, using an information-free prior (in our case, P_0^c) as the relevant vehicle (see Carnap 1971, pp. 18-19). The main message from Carnap's suggestion is the following: If the DM wishes to uncover her true belief dispositions at any point in time, then her prior probability function must be relativized only with respect to the background (non-specific) information $\mathbf{I}_B$. If the DM does so, then her current beliefs, P_m will reflect her permanent belief dispositions as well. On the contrary, if her current probabilistic beliefs are relativized to the total amount of information available at time n, namely the union of $\mathbf{I}_B$ and $\mathbf{I}_m$, then these beliefs ($P_0^{\mathbf{I}_m}$) face the risk of being emotion-laden or superficial, and hence different than P_m. For these reasons, P_0^c may alternatively be called "Carnapian prior".

At the heart of the Carnapian argument lies the view that in order for the DM to identify her true probabilistic dispositions, she must bring herself in a psychological state in which $\mathbf{I}_m$ is not treated as certainty (even if the DM actually knows $\mathbf{I}_m$), but rather as one of the many alternative, yet unrealized, possibilities. The following example lends support to the aforementioned view: Suppose that the DM, being at $t = m$, contemplates her probability of the event A : "I will live for another five years". At that time, the DM learns the information $\mathbf{I}_m$: "I am just diagnosed with lung cancer". To this end, the DM has two options: (a) The DM attempts to evaluate her subjective probability of A, under the psychological burden provoked by her viewing $\mathbf{I}_m$ as certain. In this case, she comes up with $P_0^{\mathbf{I}_m}(A) = p_1$. (b) The DM evaluates her probability of A counterfactually by asking herself the question "what would my probability of A be, were I to know that $\mathbf{I}_m$ is true?" In this case, the DM treats $\mathbf{I}_m$ as an unrealized event, which secures her a more relaxed or neutral psychological background for the evaluation of her probability of A than that of the first case. The DM's probability of A in this environment is represented by $P_0^c(A \mid \mathbf{I}_m) = p_2$. It seems reasonable to assume that $p_1 > p_2$.

Another example of how the DM's actual encounter with $\mathbf{I}_m$ might affect her ability to judge her own probabilities objectively is offered by Gul and Pesendorfer (2001): "Consider an individual who must decide what to eat for lunch. She may choose a vegetarian dish or a hamburger. In the morning, when no hunger is felt, she prefers the healthy, vegetarian dish. At lunchtime, the hungry individual experiences a craving for the hamburger." Hence, DM faces a "conflict between her ex ante ranking of options and her short-run cravings"

(2001, pp. 1403). More specifically, assume that $\mathbf{I}_m$ and A are the propositions "I am hungry" and "A vegetarian dish is conducive to longevity". If $\mathbf{I}_m$ is not realized but treated as a mere possibility then the conditional probability $P_0^c(A \mid \mathbf{I}_m)$ is (say) equal to p_2. On the contrary, if the DM calculates her own probability of A under the feeling of hunger, then $P_0^{\mathbf{I}_m}(A) = p_1$. In this case, it is quite possible that $p_2 > p_1$.

A third example of this kind comes from Greek mythology.[9] Ulysses knows already from $t = 0$ that when he will listen to sirens' song at $t = m$, he will be so enchanted by it that he will under-estimate the probability of suffering a lethal encounter with them. In an attempt to secure that at $t = m$ he will not succumb to siren's temptation, but instead he will act according to his emotionally neutral probabilistic beliefs, made at $t = 0$, the Greek hero asked his comrades to tie him up to the mast of his ship.

These examples may be thought of as a special case of a more general phenomenon pertaining to how emotional distortions impair the DM's overall ability to think objectively. Indeed, there is a plethora of empirical studies that document a negative relationship between the DM's level of anxiety (which in our case is caused by the DM's perception of $\mathbf{I}_m$ as non-contingent) and her ability to perform abstract reasoning tasks (see, for example, Leon and Revelle 1985). On another interpretation, the psychological effect of $\mathbf{I}_m$ may be thought of as a "situational moderator", which negatively affects the DM's information processing skills (see Humphreys and Revelle 1984). A common implication of both interpretations is the following: if the DM treats $\mathbf{I}_m$ as certain (that is when $P_0^{\mathbf{I}_m}(\mathbf{I}_m) = 1$), then she may experience emotional biases, which in turn impair her ability to uncover her genuine probabilistic dispositions.

5 Conclusions

In this paper, we analyzed the epistemic features of Ellsberg's "three-colors-one-urn" decision problem. This type of problems exhibits the following characteristic: The beginning of the DM's epistemic life ($t = m$) is not a time point devoid of any probabilistic information. On the contrary, at $t = m$ the DM knows that the proposition $\mathbf{I}_m$ is true (or that the event $\mathbf{I}_m$ has occurred). $\mathbf{I}_m$ is a proposition bearing "asymmetric" information, in the sense that it informs the DM about the objective probabilities of the subset of propositions $\mathcal{F}_1$ while at the same time it remains silent about the probabilities of the complement subset $\mathcal{F}_1'$ of $\mathcal{F}_1$. This means that $\mathbf{I}_m$ induces a partition $\{\mathcal{F}_1, \mathcal{F}_1'\}$ of the relevant space $\mathcal{F}$ in DM's mind. Hence, the DM enters the cognitive state of comparative ignorance, in which she feels more competent to bet on the propositions of $\mathcal{F}_1$ than on those of $\mathcal{F}_1'$. This analysis has established the following causal chain: The asymmetric information $\mathbf{I}_m$ causes the DM to feel ignorant of the events in $\mathcal{F}_1'$ compared to those in $\mathcal{F}_1$, which in turn triggers the feeling of ambiguity, thus

[9] This example is usually referred to the philosophical literature as the problem of "Ulysses and the Sirens" (see, for example, Elster 1979).

causing the DM to exhibit ambiguity aversion. A consequence of this aversion is that the DM forms incoherent beliefs and makes non-Bayesian choices.

It is important to note that the aforementioned term "ambiguity" signifies a different epistemic/cognitive state than that indicated by the term "uncertainty". What causes the DM to exhibit Ellsberg-type behavior is not uncertainty *per se*, but rather uncertainty conjoined with risk within the same decision situation. In other words, ambiguity as opposed to uncertainty is a relative concept, arising only in a comparative context such as $\{\mathcal{F}_1, \mathcal{F}_1'\}$. Therefore, if the DM is in a state of uncertainty, that is if she is uniformly ignorant about the propositions of the full set $\mathcal{F}$, then it is quite unlikely (or even unnatural) for her to form non-additive beliefs, especially in view of the fact that POI is both appealing and applicable within the "small-world" context of Ellsberg's "three-colors-one-urn" case.

The problem, however is that due to the specific feature of this decision problem mentioned above (namely the presence of $\mathbf{I}_m$), the DM does not entertain an actual time point, $t = 0$ at which she actually is in the state of uniform ignorance. As a result, it is impossible for her to develop an actual information-free prior P_0^a. This problem may be circumvented by the DM's moving into a counterfactual mode of thinking, in which she forms her prior probability function under the supposition that $\mathbf{I}_m$ is a contingent proposition, that is $P_0^c(\mathbf{I}_m) < 1$. In other words, the DM should view $\mathbf{I}_m$ not as a validated true proposition (even if it is actually such one), but rather (counterfactually) as an uncertain one on a par with any other information proposition that carries (in the DM's own standards) a non-zero probability of being true. Once this step is completed, the DM may bring $\mathbf{I}_m$ to her current system of beliefs, P_m, by Bayesian conditionalization. In such a case, P_m is additive and Ellsberg's paradox is dissolved.

Of course, whether the DM finds the aforementioned counterfactual strategy appealing is an empirical matter. Put differently, the question is whether such incoherent beliefs are "robust to clarification of the available options", where the set of options include not only actual but counterfactual ones as well. The hypothesis to be tested is whether the DM sticks to her initial non-Bayesian beliefs even after she is presented with the aforementioned counterfactual alternative of generating her prior. To that end, we may end up with one of the following cases: a) The DM finds the proposed counterfactual strategy as an attractive (though initially unconceived) alternative, thus revising the way of forming her priors accordingly, for example, by adopting the uniform prior. In this case, Ellsberg's paradox is dissolved and Bayesian rationality prevails. b) The DM is not convinced by the counterfactual suggestion, perhaps because she finds such mode of thinking unnatural. In this case, this evidence may be interpreted in two diametrically opposite ways: First, a strict Bayesian views the DM's reluctance to revise her strategy as additional evidence for her irrationality. If anything, the DM's irrationality status is elevated to that of "stubborn irrationality". Second, a more liberal Bayesian (a Bayesian with "a human face", to borrow Jeffrey's (1983) terminology) may be inclined to relax the rigid rationality criteria of strict Bayesianism, so that to accommodate Ellsberg-type choices. In this case, the theoretical literature on "ambiguity aversion" that suggests axiomatic systems of preferences different than that of Savage, becomes quite relevant. Whether the DM sticks to her initial non-Bayesian beliefs or not, once the counterfactual option has been explained to her, is an interesting empirical question that calls for careful experimental design.

References

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

Bertrand, J. (1889). Calcul des probabilites, 1st ed. (Gauthier-Villars, Paris).

Carnap, R. (1962). Logical Foundations of Probability. 2nd ed. *University of Chicago Press*.

Carnap, R. (1971). Inductive logic and rational decision. *Studies in Inductive Logic and Probability*, 1, 5-31.

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

Cheng, X. (2022). Relative maximum likelihood updating of ambiguous beliefs. J*ournal of Mathematical Economics*, 99, 102587.

Chow, C. C., and Sarin, R. K. (2001). Comparative ignorance and the Ellsberg paradox. *Journal of Risk and Uncertainty*, 22, 129-139.

Curley, S.P., Yates, J.F. and Abrams, R.A. (1986). Psychological sources of ambiguity avoidance. *Organizational behavior and human decision processes*, 38(2), 230-256.

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

Elster, J. (1979). Ulysses and the Sirens: Studies in Rationality and Irrationality. *Cambridge University Press*.

Fox, C. R. and Tversky, A. (1995). Ambiguity aversion and comparative ignorance. *The Quarterly Journal of Economics*, 110(3), 585-603.

Gilboa, I., and Marinacci, M. (2016). Ambiguity and the Bayesian paradigm. *Readings in formal epistemology*, 385-439.

Gilboa, I., and Schmeidler, D. (1993). Updating ambiguous beliefs. *Journal of Economic Theory*, 59(1), 33-49.

Gilboa, I. and Scheidler, D. (1989). Maxmin expected utility with a nonunique prior. *Journal of Mathematical Economics,* 18, 141–153.

Gul, F. and Pesendorfer, W. (2001). Temptation and self-control. *Econometrica*, 69(6), 1403-1435.

Heath, C. and Tversky, A. (1991). Preference and belief: Ambiguity and competence in choice under uncertainty. *Journal of Risk and Uncertainty*, 4(1), 5-28.

Humphreys, M. S. and Revelle, W. (1984). Personality, motivation, and performance: a theory of the relationship between individual differences and information processing. *Psychological Review*, 91(2), 153-184.

Jaynes, E. (1968). Prior probabilities. *IEEE Transactions On Systems Science and Cybernetics*, 4(3), 227–241.

Leon, M. R. and Revelle, W. (1985). Effects of anxiety on analogical reasoning: a test of three theoretical models. *Journal of Personality and Social Psychology*, 49(5), 1302-1315.

Lewis, D. (1976). Probabilities of conditionals and conditional probabilities. *Philosophical Review*, 85, 297-315.

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

Meacham, C. J. (2016). Ur-priors, conditionalization, and ur-prior conditionalization. *Ergo: An Open Access Journal of Philosophy*, 3.

Pahlke, M. (2022). Dynamic consistency in incomplete information games with multiple priors. *Games and Economic Behavior*, 133, 85-108.

Pittis, N., Koundouri, P., Samartzis, P., Englezos, N., and Papandreou, A. (2021). Ambiguity aversion, modern Bayesianism and small worlds. *Open Research Europe*, 1, 13.

Raiffa, H. (1961). Risk, Ambiguity, and the Savage Axioms: Comment. *The Quarterly Journal of Economics,* 75(4), 690-694.

Rosenkrantz, R. D. (1982). Does the philosophy of induction rest on a mistake? *The Journal of Philosophy*, 79(2), 78-97.

Savage, L. (1954). The foundations of statistics. New-York, John Wiley.

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

Skyrms, B. (1983). Zeno's paradox of measure. *Physics, Philosophy and Psychoanalysis*, 223-254.

Strevens, M. (2017). Scientific Sharing: Communism and the Social Contract. *Scientific Collaboration and Collective Knowledge*, edited by Thomas Boyer-Kassem, Conor Mayo-Wilson and Michael Weisberg. Oxford University Press.

Van Fraassen, B. C. (1989). Laws and symmetry. Clarendon Press.

6 Appendix

Proof of Proposition (3):

Let us first describe the agent's epistemic background at $t = 0$. Denote by $\mathbf{i} = (i_R, i_B, i_Y)' \in \boldsymbol{I}$ the 3×1 vector that contains the numbers of red, black and yellow balls in the urn, respectively, where $\boldsymbol{I} = \{\mathbf{i} \in \mathbb{N}^3 : \mathbf{0} \leq \mathbf{i} \leq \mathbf{1}{\times}90$ and $\mathbf{1}'{\times}\mathbf{i} = 90\} \subset \mathbb{N}^3$, and $\mathbf{1} = (1,1,1)'$. For convenience, we also define $\boldsymbol{I}_a^k \subset \boldsymbol{I}$, to be the subset of vectors for which the number of $a-$colored balls in the urn, is exactly k, where $a \in \{R, B, Y\}$, and $0 \leq k \leq 90$.

First of all, the agent has to decide about her prior probabilities of the hypotheses in $\mathbf{H}$. The agent, having no reason to consider one proposition more likely than another, adopts the principle of indifference, which for the present case (in which the number of propositions is finite) is identical to both Leibnitz's "principle of insufficient reason" and Jaynes' "principle of maximum entropy" (Jaynes 1968). Therefore, she equates equal probabilities among $\mathcal{H}_{\mathbf{i}} \in \mathbf{H}$ and in particular,

$$P_0^c(\mathcal{H}_{\mathbf{i}}) = \frac{1}{4,186}, \mathbf{i} \in \boldsymbol{I}.$$

The important thing to notice is that there is no specific information at $t = 0$, hence there is no informational asymmetry between the hypotheses $\mathcal{H}_{\mathbf{i}}, \mathbf{i} \in \boldsymbol{I}$.

Using the law of total probability, the agent gets:

$$P_0^c(s_a) = \sum_{\mathbf{i} \in \boldsymbol{I}} P_0^c(s_a \mid \mathcal{H}_{\mathbf{i}}) P_0(\mathcal{H}_{\mathbf{i}}).$$

It is easy to show that, $P_0^c(s_a \mid \mathcal{H}_{\mathbf{i}}) = \frac{k}{90}$, if $\mathbf{i} \in \boldsymbol{I}_a^k$ and $card(\boldsymbol{I}_a^k) = 91 - k$, $\forall k = 0, ..., 90$ and $\forall a \in \{R, B, Y\}$. As a result,

$$\begin{aligned} P_0^c(s_a) &= \sum_{\mathbf{i} \in \boldsymbol{I}} P_0^c(s_a \mid \mathcal{H}_{\mathbf{i}}) P_0(\mathcal{H}_{\mathbf{i}}) = \\ &= \sum_{k=0}^{90} \sum_{\mathbf{i} \in \boldsymbol{I}_a^k} P_0^c(s_a \mid \mathcal{H}_{\mathbf{i}}) P_0(\mathcal{H}_{\mathbf{i}}) = \\ &= \sum_{k=0}^{90} \frac{k}{90} (91 - k) \frac{1}{4,186} = \frac{1}{3}, \forall c \in \{R, B, Y\}. \end{aligned}$$

Therefore,

$$P_0^c(s_R) = P_0^c(s_B) = \frac{1}{3} \text{ and } P_0^c(s_{RY}) = P_0^c(s_{BY}) = \frac{2}{3}.$$

Clearly, the agent will be indifferent between actions f and g and between actions f^* and g^* in the absence of any specific information.

At time $t = 1$ the agent acquires an important piece of specific information for the problem at hand. In particular she is given the information that the number of red balls in the urn is l, i.e. she finds out $I_S = \boldsymbol{I}_R^l =$ "the urn contains l red balls", where $0 \leq l \leq 90$. Note that in the standard version of Ellsberg paradox, $l = 30$.

Bayesian conditionalization implies that $\forall a \in \{R, B, Y\}$:

$$P_1(s_a) = P_0^c\left(s_a \mid I_S\right) = \sum_{\mathbf{i} \in \boldsymbol{I}} P_0^c\left(s_a \mid \mathcal{H}_{\mathbf{i}} \wedge I_S\right) P_0^c\left(\mathcal{H}_{\mathbf{i}} \mid I_S\right),$$

where

$$P_0^c\left(s_a \mid \mathcal{H}_{\mathbf{i}} \wedge I_S\right) = \begin{cases} P_0^c(s_a \mid \mathcal{H}_{\mathbf{i}}), & \mathbf{i} \in \boldsymbol{I}_R^l \\ 0, & \mathbf{i} \notin \boldsymbol{I}_R^l \end{cases}$$

and

$$P_0^c\left(\mathcal{H}_{\mathbf{i}} \mid I_S\right) = \begin{cases} \frac{1}{91-l}, & \mathbf{i} \in \boldsymbol{I}_R^l \\ 0, & \mathbf{i} \notin \boldsymbol{I}_R^l \end{cases}.$$

Therefore,

$$P_1(s_a) = \sum_{\mathbf{i} \in \boldsymbol{I}} P_0^c\left(s_a \mid \mathcal{H}_{\mathbf{i}} \wedge I_S\right) P_0^c\left(\mathcal{H}_{\mathbf{i}} \mid I_S\right) = \sum_{\mathbf{i} \in \boldsymbol{I}_R^l} P_0^c(s_a \mid \mathcal{H}_{\mathbf{i}}) \frac{1}{91-l}, \forall a \in \{R, B, Y\}.$$

Clearly,

$$P_1(s_R) = \sum_{\mathbf{i} \in \boldsymbol{I}_R^l} P_0^c(s_R \mid \mathcal{H}_{\mathbf{i}}) \frac{1}{91-l} = \sum_{k=0}^{90-l} \frac{l}{90} \frac{1}{91-l} = \frac{l}{90}.$$

Moreover, $P_0^c(s_B \mid \mathcal{H}_{\mathbf{i}}) = \frac{k}{90}$, if $\mathbf{i} \in \boldsymbol{I}_R^l \cap \boldsymbol{I}_B^k$, $P_0^c(s_Y \mid \mathcal{H}_{\mathbf{i}}) = \frac{k}{90}$, if $\mathbf{i} \in \boldsymbol{I}_R^l \cap \boldsymbol{I}_Y^k$ and $card(\boldsymbol{I}_R^l \cap \boldsymbol{I}_B^k) = card(\boldsymbol{I}_R^l \cap \boldsymbol{I}_Y^k) = 1$, $\forall k = 0, ..., 90 - l$. As a result,

$$P_1(s_a) = \sum_{k=0}^{90-l} \sum_{\mathbf{i} \in \boldsymbol{I}_R^l \cap \boldsymbol{I}_a^k} P_0^c(s_a \mid \mathcal{H}_{\mathbf{i}}) \frac{1}{91-l} = \sum_{k=0}^{90-l} \frac{k}{90} \frac{1}{91-l} = \frac{90-l}{180}, a \in \{B, Y\}.$$

Finally,

$$P_1(s_R) = \frac{l}{90}, P_1(s_B) = \frac{90-l}{180} \text{ and } P_1(s_{RY}) = \frac{90+l}{180}, P_1(s_{BY}) = \frac{90-l}{90}.$$

From the last equations it follows that P_1 defined on $\mathcal{F}$ is additive (and therefore adequate).

The previous analysis shows that $P_1(s_B) = P_1(s_Y) = \frac{90-l}{180}$, i.e. the agent, at time t, is indifferent between the propositions for which she has no specific information. A question that naturally arises, is whether this indifference is the reason why there is no contradiction. To see whether this is the case, we assume that

$$P_0^c(\mathcal{H}_{\mathbf{i}}) = p_{\mathbf{i}},$$

where $p_{\mathbf{i}} > 0$ and $\sum_{\mathbf{i} \in \boldsymbol{I}} p_{\mathbf{i}} = 1$. In this case,

$$P_0^c\left(\mathcal{H}_{\mathbf{i}} \mid I_S\right) = \begin{cases} \frac{p_{\mathbf{i}}}{\sum_{\mathbf{j} \in \boldsymbol{I}_R^l} p_{\mathbf{j}}}, & \mathbf{i} \in \boldsymbol{I}_R^l \\ 0, & \mathbf{i} \notin \boldsymbol{I}_R^l \end{cases}.$$

Therefore,

$$P_1(s_a) = \sum_{k=0}^{90-l} \sum_{\mathbf{i} \in \boldsymbol{I}^l_R \cap \boldsymbol{I}^k_a} P^c_0(s_a \mid \mathcal{H}_{\mathbf{i}}) P^c_0(\mathcal{H}_{\mathbf{i}} \mid I_S) =$$

$$= \sum_{k=0}^{90-l} \frac{k}{90} \sum_{\mathbf{i} \in \boldsymbol{I}^l_R \cap \boldsymbol{I}^k_a} \frac{p_{\mathbf{i}}}{\sum_{\mathbf{j} \in \boldsymbol{I}^l_R} p_{\mathbf{j}}}, a \in \{B, Y\}.$$

As a result,

$$P_1(s_R) = \frac{l}{90},$$

$$P_1(s_B) = \sum_{k=0}^{90-l} \frac{k}{90} \sum_{\mathbf{i} \in \boldsymbol{I}^l_R \cap \boldsymbol{I}^k_B} \frac{p_{\mathbf{i}}}{\sum_{\mathbf{j} \in \boldsymbol{I}^l_R} p_{\mathbf{j}}} = \frac{E_0(s_B \mid I_S)}{90} \text{ and}$$

$$P_1(s_{RY}) = \frac{l}{90} + \sum_{k=0}^{90-l} \frac{k}{90} \sum_{\mathbf{i} \in \boldsymbol{I}^l_R \cap \boldsymbol{I}^k_Y} \frac{p_{\mathbf{i}}}{\sum_{\mathbf{j} \in \boldsymbol{I}^l_R} p_{\mathbf{j}}} = \frac{l + E_0(s_Y \mid I_S)}{90},$$

$$P_1(s_{BY}) = \frac{(90 - l)}{90}.$$

Again, P_1 defined on $\mathcal{F}$ is additive (and therefore adequate).