DEPARTMENT OF INTERNATIONAL AND EUROPEAN ECONOMIC STUDIES

ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

Mitigating Ambiguity Aversion via Counterfactual Priors: A Resolution of Ellsberg's Paradox

Phoebe Koundouri

Nikitas Pittis

Panagiotis Samartzis

Working Paper Series

22-13

March 2022

Mitigating Ambiguity Aversion via Counterfactual Priors: A Resolution of Ellsberg's Paradox

Phoebe Koundouri*† Nikitas Pittis‡ Panagiotis Samartzis*§

Abstract

Ellsberg-type preferences violate one of the principles for Bayesian rationality, namely Savage's Sure Thing Principle, and are among the main empirical results against Subjective Expected Utility theory. In this paper, we propose a novel strategy for dealing with ambiguity aversion and the resulting Ellsberg-type choices. First, we identify the presence of "asymmetric information" as the main cause of ambiguity aversion. Second, we develop a solution for Ellsberg's paradox which emerges as a direct application of counterfactual thinking, implemented to the specific choice problem described by Ellsberg. Third we analyze the psychological, methodological and logical merits of the developed counterfactual strategy, and show that its application solves the problems of "error correction" and "unconceived alternatives", two of the main complaints about Bayesian Confirmation Theory.[1]

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 Banking and Financial Management, University of Piraeus, Greece; School of Economics, Athens University of Economics and Business, Greece; International Center for Research on the Environment and the Economy, Greece; e-mail: psamartzis@unipi.gr

1 Introduction

It is widely acknowledged that the so-called Ellsberg's paradox poses one of the most serious threats against the empirical validity of the "Subjective Expected Utility Maximization" (SEUM) theory. It has long been recognized that such choices violate one of Savage's (1954) axioms for rational preferences, namely his illustrious Sure Thing Principle (STP). As a consequence, the decision maker (DM) who exhibits Ellsberg-type choices is not probabilistically sophisticated and does not adhere to SEUM criterion of choice. This means that the degrees of her probabilistic beliefs (or credences) are not *coherent*, that is they are not represented by a unique subjective probability function that obeys the axioms of Kolmogorov (one of which is the axiom of additivity - finite or countable). Put differently, such a DM is unable to come up with a unique numerical probability for each and every event/proposition of the relevant space.

What is the "degree of rationality" of such a probabilistically non-sophisticated / non SEU-maximizer DM? Should be DM condemned as "non-rational" when she exhibits Ellsberg-type behaviour? Or alternatively, the requirement of probabilistic sophistication is too strict to be treated as part of the relevant normative ideal? Is it possible that DM's inability to express probabilistically sophisticated beliefs is (in some cases) justified, and as such, it should not be interpreted as evidence against DM's rationality?

One interesting case, pertaining to the aforementioned distinction is that of "ambiguity aversion" (AA). Roughly, a DM is ambiguity averse if she prefers to bet on events/propositions with known probabilities than on those with unknown ones. Is such a DM rational? In other words, does AA reflect a pathological choice of a non-rational DM, or should AA be included in DM's normative system of preferences? In the latter case (as opposed to the former one), AA may be thought of as an acceptable property of DM's preferences, on a par with other such properties, such as completeness and transitivity. Many authors adopt this more tolerant definition of rationality and proceed in developing a system of preferences in which AA plays a prominent role. Such axiomatic systems have been proposed (among many others) by Schmeidler (1989) and Gilboa and Schmidler (1989). A common element of these systems is the abolition of STP and the addition of alternative axioms which state explicitly what it means for DM's preferences to display ambiguity aversion. Each of these systems explains Ellsberg's paradox in the sense that within the system, Ellsberg's choices become predictable. Moreover, each system spells out the type of (non-sophisticated) probabilistic beliefs that an ambiguity averse DM possesses. In Schmeidler (1989) and Gilboa and Schmeidler (1989), for example, these beliefs are represented by a convex (non-additive) probability function and by multiple probability functions, respectively. It is important to note that although such probabilistic beliefs are technically referred to as non-sophisticated, they are not irrational. Indeed, as Levi (1985) puts it, DM's refusal to make a determinate probability judgment "may derive from a very clear and cool judgment that on the basis of the available evidence, making a numerically determinate judgment would be unwarranted and arbitrary." (1985, pp. 396).

Contrary to the aforementioned interpretation, another strand of the literature (mostly strict Bayesians) views AA as unambiguous evidence of non-rationality (see, for example, Raiffa 1961, Savage 1974). This literature identifies the origins of AA, not in DM's careful deliberation on the co-existence of clear and vague alternatives, but rather in her instinctive psychological reaction towards avoiding the ambiguous option: "In expressing a preference for the less ambiguous option, subjects are making a nonconscious, systematic error which, if sufficiently understood, they would correct." (Curley et. al 1986, pp. 233). This part of the literature views AA as a pathological symptom of DM's psychology, and as such it cannot be reconciled with DM's rationality. Why is AA supposed to be inconsistent with rationality? The authors who express such an opinion claim that DM's ambiguity aversion along with the resulting incoherence of her subjective probability function make her vulnerable to a Dutch book. The latter is a series of bets (offered to DM by a cunning bettor, not better informed than DM) that DM is willing to accept individually, but which jointly inflict upon her a sure loss. A rational DM, so the argument goes, would never accept a set of bets that deterministically lead to her losing money. Hence, if DM does so, she is not rational and needs probabilistic education.

What is the cause of AA? One of the most prominent psychological explanations for AA is the so-called "comparative ignorance hypothesis", put forward by Heath and Tversky (1991) and Fox and Tversky (1995). The main tenet of this hypothesis is that AA is caused by DM's spontaneous disposition towards preferring acts about which she feels to be more knowledgeable, rather than those for which she feels to be ignorant: "Thus, ambiguity aversion represents a reluctance to act on inferior knowledge, and this inferiority is brought to mind only through a comparison with superior knowledge about other domains or of other people." (Fox and Tversky 1995 pp. 599). What generates the distinction between "superior" and "inferior" knowledge? We argue that this distinction is caused by asymmetric information about DM's events of interest. One type of asymmetric information is *probabilistic asymmetric information* and this is the kind of asymmetry that gives rise to Ellsberg-type choices. In particular, the specific information, I_S, that DM possesses induces a partition $\{\mathcal{F}_k, \mathcal{F}'_k\}$ of the space of events/propositions $\mathcal{F}$, with $\mathcal{F}_k$ and $\mathcal{F}'_k$ being the subspaces with the known and unknown objective probabilities, respectively. [2]

The aforementioned problem of DM's comparative ignorance may be reiterated in terms of the manner in which DM chooses to process the specific information I_S. We must first recognize that Ellsberg-type choices correspond to a specific epistemic state for DM: Her epistemic life, that is her deliberations about the phenomenon of interest, begins *at the same time*, say $t = 1$, that she learns about the truth of I_S. This means that any time point prior to $t = 1$, for example $t = 0$, is devoid of any epistemic activity on the part of DM, including

[2] There are other types of asymmetric information that may cause AA which are not probabilistic. For example, the aforementioned authors show that DM is likely to experience comparative ignorance if she feels to be more knowledgeable about the events of $\mathcal{F}_k$ than those of $\mathcal{F}'_k$.

her constructing a prior probability function P_0. As a result, the time point $t = 1$ finds DM "probabilistically unprepared", in the sense that she does not possess a pre-existing prior probability function P_0, which she could have used as a vehicle for updating her prior beliefs in the light of I_S at $t = 1$. Hence, at $t = 1$, DM has to accomplish two epistemic tasks simultaneously: First to form (for the first time) her probabilistic beliefs, P_1, about the elements of $\mathcal{F}$ and second to process the information content of I_S. These two tasks, however, are not independent but instead they interact with each other. Comparative ignorance may be thought of as the result of this interaction, that is DM's attempt to come up with her probabilistic assessments for the elements of $\mathcal{F}$ under the influence of her knowing that I_S is true. The result of DM's attempt to accomplish the aforementioned dual task is to form incoherent degrees of belief, hereafter referred to as $P_1^{I_S}$.

The last point generates the following question: In the presence of the specific asymmetric information I_S, what can DM do in order to avoid forming probabilistic beliefs under the state of comparative ignorance? One possible strategy for DM would be to seek the "missing information", that is to find the specific probabilistic information $I_{S'}$ pertaining to $\mathcal{F}_k'$. If such an option is feasible, then $I_S \cup I_{S'}$ is no longer asymmetric and DM will effectively move from the state of ambiguity to that of risk. If, however, the acquisition of $I_{S'}$ is not possible, then DM is left with only one alternative option that may drive her out of ambiguity, namely the following two-step counterfactual strategy: In the first step, DM temporarily "deletes" I_S from her total corpus of available information and evaluates her subjective probability function by asking herself the following question: "what would my subjective probability function P_0^c be, had I not known that I_S is true?" By making this counterfactual move, DM brings herself in an epistemic state in which she is unsure of the truth of I_S. In other words, DM's turn to counterfactual thinking aims at changing temporarily the modal status of I_S from "certainty" to "possibility". In the next sections, we argue that this specific epistemic state, in which I_S is not treated by DM as certainty, but as one of the many yet unrealized possibilities, or as one of the many alternative pieces of information that DM will eventually receive in the "future", is the only legitimate state for DM to identify her genuine probabilistic beliefs. Once the aforementioned counterfactual question is answered and DM's prior "I_S-free" probability function P_0^c is built, DM proceeds to the second step in which she brings I_S back to the picture by updating her prior beliefs via Bayesian conditionalization on I_S, using P_0^c as the relevant vehicle of conditionalization. These two steps form the proposed two-step, indirect counterfactual strategy, hereafter referred to as IC_{I_S}. The key element of IC_{I_S} is that in forming her probabilistic beliefs of $\mathcal{F}$, DM has (counterfactually) brought herself in the epistemic state of *uniform* rather than *comparative* ignorance, in which DM is not susceptible to ambiguity aversion. Note that the term "indirect" in IC_{I_S} refers to the manner in which DM chooses to process I_S : Instead of allowing I_S to directly affect her probabilistic beliefs at the time of their formation, DM permits I_S to affect them indirectly via conditionalization, that is $P_1(A) = P_0^c(A \mid I_S)$, $A \in \mathcal{F}$ with $P_0^c(I_S) < 1$.

The rest of the paper is organized as follows: Section 2 analyzes the psychological causes of AA that have been proposed in the literature, paying special emphasis to the one that best captures the spirit of Ellsberg's paradox, namely comparative ignorance caused by asymmetric probabilistic information. It also shows how the proposed counterfactual strategy IC_{I_S} may drive DM out of comparative ignorance, thus eliminating AA and restoring DM's probabilistic sophistication. Section 3 presents several psychological arguments supporting the formation of counterfactual priors. We also argue that the merits of counterfactual priors were already identified in 1961 by the great philosopher of science Rudolph Carnap. Section 4 contains the main results of the paper. In particular, it revisits Ellsberg's paradox and demonstrates how this paradox is resolved by means of the proposed counterfactual strategy IC_{I_S}.

The rest of the paper attempts to answer the following question: "Is IC_{I_S} a procedure especially designed to solve Ellsberg's paradox? Or is it possible that IC_{I_S} exhibits some virtues that extend well beyond the narrow scope of solving this particular paradox?" To this end, we distinguish two types of reasons that make IC_{I_S} attractive, namely methodological and logical. Section 5 presents our methodological arguments in favor of IC_{I_S} by showing how the adoption of this strategy addresses two of the main pitfalls of the Bayesian Confirmation Theory, namely the problems of "error correction" and "unconceived alternatives". In Section 6, we offer our logical arguments for IC_{I_S} by analyzing the logical structure of the concept of "information processing" and showing that counterfactual probabilistic reasoning is an indispensable part of it. Section 7 concludes the paper.

2 Psychological Origins of Ambiguity Aversion: Epistemic Reliability and Second-order Probabilities

This section analyzes the potential sources of AA. In particular, it focuses on the characteristics of DM's epistemic conditions that drive her into the psychological state of ambiguity.

Using the Knightian distinction between "risk" and "uncertainty", DM faces a risky situation (known probabilities supplied by I_S) with respect to $\mathcal{F}_k$, whereas she operates in an environment of uncertainty (unknown probabilities) with respect to $\mathcal{F}'_k$. Should DM distinguish between risk and uncertainty and especially should she prefer the former over the latter? According to strict Bayesians, the answer is negative: DM is always able to ascertain her own subjective probabilities of $\mathcal{F}'_k$ which in combination with the known probabilities of $\mathcal{F}_k$ yield a proper subjective probability function over the whole of $\mathcal{F}$. In the same spirit, Levi (1985) remarks: "A rational agent is committed to recognizing a single probability function for use in computing expected utilities in any given context of deliberation" (1985, pp. 391). [3] This means that AA sets in when

[3] In fact in some forms of radical subjectivism, DM is allowed to ignore the furnished

DM does not treat risk and uncertainty symmetrically; instead she prefers the former epistemic state over the latter.

Why does DM exhibit such a non-Bayesian preference pattern? One answer to this question may be given by the comparative ignorance hypothesis. Comparative ignorance arises when DM believes that the degree of epistemic reliability that she assigns to the known probabilities of $\mathcal{F}_k$ (furnished by I_S) is much greater than the one she assigns to the unknown probabilities of $\mathcal{F}'_k$. Put differently, comparative ignorance demands DM to trust the exogenously given probabilities of $\mathcal{F}_k$ to a greater extent than her own probabilistic judgments of $\mathcal{F}'_k$. This difference in reliability may be expressed as follows: Let I_S be the proposition: "There are 30 red balls in the Ellsberg urn", with the latter being an urn that contains 90 red, black and yellow balls in unknown proportions. Let us further assume that DM fully trusts the source that provides her with I_S. This latter assumption is translated into $P_1(I_S) = 1$, that is her subjective probability of the truth of I_S is equal to unity. Let $A \in \mathcal{F}_k$. Based on I_S and under the additional assumption that DM obeys the Principal Principle, DM's subjective probability of A is equal to

$$P_1(A \mid I_S) = p_A.$$

For example, if A is the proposition "the next draw is a red ball", $p_A = 1/3$. How firmly does DM believe the proposition "$P_1(A \mid I_S) = p_A$"? This question brings us to the realm of "probabilities of probabilities", or second-order probabilities. If DM fully trusts I_S, in the sense $P_1(I_S) = 1$, then her subjective (second-order) probability, Q_1, for "$P_1(A \mid I_S) = p_A$" at $t = 1$ is equal to one. Specifically,

$$Q_1 \left(P_1(A \mid I_S) = p_A \mid I_S\right) = 1,\ A \in \mathcal{F}_k. \tag{1}$$

Note that DM conditionalizes on the specific information I_S twice. First, with respect to P_1, so that $P_1(A \mid I_S) = p_A$ and second with respect to Q_1, so that $Q_1 \left(P_1(A \mid I_S) = p_A \mid I_S\right) = 1$. The last equation allows us to say that all the propositions in $\mathcal{F}_k$ are characterized by the same degree of epistemic reliability, with this degree being equal to unity.

Next, let $B \in \mathcal{F}'_k$ for which no specific information is available. For example, B is the proposition "the next draw is a black ball". What is DM's subjective probability of B? There are two answers to this question, depending on whether DM is a strict Bayesian or not. First, a strict Bayesian is not affected by the fact that there is no information on B. She is fully confident in her own ability to judge the likelihood of B. Hence, she ends up with

$$P_1(B) = p_B.$$

with

$$Q_1 \left(P_1(B) = p_B\right) = 1,\ B \in \mathcal{F}'_k. \tag{2}$$

objective probabilities for $\mathcal{F}_k$, stick to her own probabilistic judgments for $\mathcal{F}_k$, and still be rational. Of course, such a DM violates the Probabilities Coordination Principle, one form of which is Principal Principle.

For a less dedicated Bayesian (moderate Bayesian), however, the situation is different. Such a DM is affected by the fact that she is less informed for $B \in \mathcal{F}_k'$ than for $A \in \mathcal{F}_k$. She argues that beliefs need to be justified, and such a justification is missing in the case of $B \in \mathcal{F}_k'$ as opposed to that of $A \in \mathcal{F}_k$. Hence, she finds a claim such as (2) absurd. Instead, she thinks that her subjective probability, $P_1(B)$, of B can take on any of the following n values, $p_{B,1}$, $p_{B,2}, ..., p_{B,n}$. How "probable" does she find each of these values? This question is answered by the second-order probabilities, $q_{B,1}$, $q_{B,2}, ..., q_{B,n}$, corresponding to $p_{B,1}$, $p_{B,2}, ..., p_{B,n}$, respectively, with $\sum_{i=1}^{n} q_{B,i} = 1$. Specifically,

$$
\begin{aligned}
Q_1 \left(P_1(B) = p_{B,1} \right) &= q_{B,1} \\
Q_1 \left(P_1(B) = p_{B,2} \right) &= q_{B,2} \\
&. \\
&. \\
&. \\
Q_1 \left(P_1(B) = p_{B,n} \right) &= q_{B,n}.
\end{aligned} \tag{3}
$$

For example, if DM is informed that "the coin is fair" then she thinks that she is justified to believe that her probability of "heads" is equal to 1/2, with (second-order) probability one. If, however, such information is missing, then she might find two probabilities of "heads" plausible, for example 0.5 or 0.6. Now, if she believes that her first probabilistic assignment is more probable than the second, she may express this additional belief by setting her second-order probabilities equal to (say) 0.9 and 0.1, respectively.

The strict Bayesian may reply that the aforementioned second-order probabilities are redundant (or even meaningless), in the sense that they do not enhance DM's credal structure in any meaningful way. This is due to the following two reasons: First, DM can always calculate a single probability of B, simply by taking the average of first-order probabilities weighted by the corresponding second-order probabilities,

$$\overline{P}_1(B) = \sum_{i=1}^{n} q_{B,i} p_{B,i}. \tag{4}$$

The last relationship shows that DM can, in principle, calculate a single probability for both $A \in \mathcal{F}_k$ and $B \in \mathcal{F}_k'$, namely $P_1(A \mid I_S)$ and $\overline{P}_1(B)$, respectively.

Second, there is a considerable difference between the objects on which first-order probabilities are attached, e.g. the proposition B and those on which second-order probabilities are attached, e.g. the proposition "$P_1(B) = p_{B,1}$". The basic difference is that B represents a proposition that assumes truth values (it can be true or false) or, equivalently, an empirical event, whereas "$P_1(B) = p_{B,1}$" does not. This is due to the fact that a proposition assumes truth values if and only if its veracity can be established by comparison with reality (or at least, experience). This is the case only if the proposition is factual. For example, the proposition "the next draw is red" is either true or false. Why is that? Because whether this proposition is true or false can be decided by observing the outcome of the next draw. If the next draw is actually red, then B "corresponds to reality" and, hence, it is true. If it does not, then it is false. The truth or falsity of B is decidable. On the contrary, the truth of the proposition "DM's probability of B is equal to 0.4" cannot be decided by means of an empirical procedure. No experiment that allows comparison of this statement with objective reality is feasible. Hence, "$P_1(B) = p_{B,1}$" cannot be either true or false. But if a "proposition" does not assume truth values, then it cannot carry a probability. The probability of a proposition is the probability that this proposition is true or, equivalently, the probability of an event is the probability that the event occurs. These remarks lead to the following conclusion. The first-order probability $P_1(B)$ is meaningful, since it represents "the probability that B is true" or "the probability that B occurs". On the other hand, the second-order probability $Q_1(P_1(B) = p_{B,1})$ cannot be interpreted as the "the probability that $P_1(B) = p_{B,1}$ is true" since the argument in the $Q(\cdot)$ function (as opposed to that in the $P(\cdot)$ function) does not assume truth values (in the sense analyzed above).

The foregoing discussion suggests that the strict Bayesian insists that a rational DM should always be able to form coherent probabilistic beliefs even in the absence of any specific information. The strict Bayesian emphasizes that the essential requirement of rationality is that DM starts her epistemic life with an intrinsically consistent set of beliefs. Coherence is the minimal requirement for rationality. On the other hand, rationality does *not* require DM's subjective probabilities to be equal to the corresponding objective ones, at least at the early phase of DM's epistemic life. If specific information about objective probabilities is missing, then there is no Bayesian demand for DM to form accurate or realistic probabilistic beliefs. At the beginning of DM's epistemic life (that is at $t = 0$), all that Bayesianism requires is coherence not realism. Realism will come later with the aid of empirical data and using conditionalization as the proper learning protocol. Hence, all that the strict Bayesian requires is with respect to $\mathcal{F}'_k$ is that DM has coherent probabilistic beliefs, regardless of whether these beliefs are inaccurate with respect to objective chances. The strict Bayesian has no problem of accepting that $\mathcal{F}_k$ and $\mathcal{F}'_k$ are, indeed, characterized by different degrees of epistemic reliability. She is also probably willing to consent that DM is comparatively more knowledgeable for $\mathcal{F}_k$ than $\mathcal{F}'_k$. She takes issue however, with what epistemic reliability and comparative ignorance are supposed to entail, namely ambiguity aversion and non-sophisticated probabilistic beliefs. She emphasizes that comparative ignorance may affect the realism of probabilistic beliefs but not their coherence.

Despite the normative force of the strict Bayesian's arguments, a moderate Bayesian finds comparative ignorance to be a perfectly good reason (or excuse) for forming incoherent beliefs. She would also add that what makes her incoherent is not the mere fact that the probability of $A \in \mathcal{F}_k$ is single-valued, whereas the probability of $B \in \mathcal{F}'_k$ is multi-valued. Instead, the origins of the problem are traced somewhere else, namely in that DM's probability of A takes on a single value *with certainty*, whereas her probability of B is itself a *random variable,* whose values are DM's first-order probabilities. When the two options A and B are offered to DM *simultaneously*, that is when DM is given the opportunity to *compare* the two options, she is inclined to choose the certain one, that is the option with the maximum degree of epistemic reliability. Needless to say, a moderate Bayesian finds such a behaviour to be quite rational.

We have been led to the following conclusions: When DM has to choose among bets of the same kind, that is among propositions with the same degree of epistemic reliability, DM is not psychologically inclined to favor some propositions over the others. She treats them all symmetrically. The Ellsberg-type choices arise only when DM has to choose among a set of propositions with different degrees of epistemic reliability. In other words, DM exhibits AA only when she contrasts options of different kinds (or options belonging to different classes), for example propositions with known probabilities to propositions whose probabilities are random variables. Hence, the moderate Bayesian is likely to consent that if she were in the epistemic state of uniform rather than comparative ignorance, that is if she were equally uninformed about $\mathcal{F}_k$ and $\mathcal{F}'_k$, then she would exhibit no ambiguity aversion.

The preceding discussion highlights the instrumental role of the asymmetric specific information I_S for AA. To this end, the causal chain of events that leads to AA is the following: First, the specific information I_S, which informs DM about the probabilities of $\mathcal{F}_k$ alone, becomes available at $t = 1$. Second, DM's confidence on the truth of I_S is absolute, that is $P_1(I_S) = 1$. As a result, her degree of epistemic reliability for the probabilities of $\mathcal{F}_k$, furnished by I_S, is maximum, as stated by (1). On the other hand, DM feels much less confident about her own probabilistic judgments for the elements of $\mathcal{F}'_k$, as stated by (3). At this point, DM experiences the feeling of comparative ignorance. Third, as a reaction to comparative ignorance, DM exhibits AA.

The preceding discussion reveals what the proposed counterfactual strategy IC_{I_S} is designed to accomplish: It aims at bringing DM in an epistemic state in which she does not (actually) believe that I_S is true with probability one, but rather she thinks of it (counterfactually) as one of many alternative eventualities that may come to be true "in the future". Specifically, IC_{I_S} aims at replacing $P_1(I_S) = 1$ with $P^c_0(I_S) = p_0 < 1$ in DM's mind. By making this counterfactual move, DM brings herself in an epistemic state in which she is equally uninformed about the probabilities of $\mathcal{F}_k$ and $\mathcal{F}'_k$, that is she restores her "uniform ignorance" over $\mathcal{F}$. In this state, all her probabilistic assignments on $\mathcal{F}$ exhibit the same degree of epistemic reliability, and because of this, DM is guarded against entering the cognitive state of comparative ignorance.

It must be emphasized at this point that DM's shift from the epistemic state of comparative ignorance to that of uniform ignorance aims at achieving one task only: To drive DM out of the comparative mode of probabilistic thinking, that is to make her think not in relative terms ($\mathcal{F}_k$ relative to $\mathcal{F}'_k$) but in absolute terms (treating the elements of $\mathcal{F}$ uniformly). Is it certain that DM will prove to be probabilistically sophisticated in the new epistemic environment of uniform ignorance? Definitely not. In the absence of any specific information, DM finds herself in the epistemic state of Knightian uncertainty and because of this, she may find it hard to come up with a precise numerical assignment for each proposition of $\mathcal{F}$. However, this type of probabilistic predicament is a general challenge for Bayesianism and not the reason that gives rise to Ellsberg's paradox. The type of uncertainty that characterizes Ellsberg's paradox is a special type of uncertainty (referred to as ambiguity) arising from DM's simultaneous comparison of bets with known probabilities to bets with unknown ones. IC_{I_S} aims at driving DM out of ambiguity and bringing her back in the epistemic context of classical uncertainty. What are the gains from such a maneuver? DM is back to the uncertainty environment, which is encoded by Savage's axioms. In the absence of comparative ignorance, DM has one less reason to violate STP. This implies that the probability of DM being probabilistically coherent is higher in the epistemic state of uncertainty, than it is in the state of ambiguity. Put differently, a DM who under uncertainty has succeeded in forming coherent probabilistic beliefs may fail to do so if she moves to the state of ambiguity. Comparative ignorance is an additional reason for causing incoherence, over and above the reasons that are present under the state of uniform ignorance.

To further elaborate the point made in the last paragraph, let us consider a DM who is about to draw a ball from Ellsberg's urn. Let us compare the way she forms her subjective probabilities under uncertainty, to that under ambiguity. Under uncertainty, she has no information about the relative frequencies of the three colors. This knowledge environment provides DM with a "natural symmetry", namely three possible colours for which no specific information is available. As a result, DM is likely to find Laplace's "principle of indifference" quite fitting, thus assigning equal probabilities to the three colors, namely $P(R) = P(Y) = P(B) = 1/3$. Moreover, DM's degree of epistemic reliability for these three probabilities is the same, which in turn implies that DM's betting propensity is uniformly distributed among the three colors. Under ambiguity, however, the situation is likely to change drastically. Now, DM is assumed to possess the information I_S that 30 of the 90 balls are red. Hence, she sets her subjective probability, $P^{I_S}(R)$, of "red" equal to the corresponding objective probability, namely 1/3. What is the difference between DM's probability of "red" under uncertainty and the corresponding probability under ambiguity, that is between $P(R)$ and $P^{I_S}(R)$? Note, that these two probabilities are both equal to 1/3. Does this mean that DM's credal state about "red" under uncertainty is identical to that under ambiguity? The answer is negative. DM has a good reason to trust the proposition $P^{I_S}(R) = 1/3$ much more than the proposition $P(R) = 1/3$, with this reason being no other than the presence of the reliable I_S. Despite their equality, $P(R)$ and $P^{I_S}(R)$ do not reflect identical probabilistic attitudes for "red". How does DM's increased confidence for "red" affect her probabilistic assignments for "yellow" and "black"? One might argue that DM is still free to set $P^{I_S}(Y) = P^{I_S}(B) = 1/3$. However, if she does so her betting behaviour towards the three colors has to be symmetric, that is she must not prefer to bet on any one color over another. As already mentioned, there is significant empirical evidence against the aforementioned symmetry hypothesis; people tend to prefer to bet on "red" rather than "yellow" or "black". In such a case, DM's implicit probabilities for "yellow" and "black" do not cohere with $P^{I_S}(R) = 1/3$. Hence, DM ends up being probabilistically incoherent. To conclude, the arrival of I_S brings about two distinct and interdependent effects on DM's credal state: On the one hand, it increases DM's corpus of knowledge about the objective probabilities of the events of interest. On the other hand, it demolishes the "natural symmetry" feature of the phenomenon at hand, thus causing DM's incoherence. The question which naturally arises at this point is the following: How should DM process I_S so that she combines the best of both worlds, namely increased knowledge of objective probabilities with coherence of subjective ones? As already argued, IC_{I_S} is an information-processing strategy, precisely designed to accomplish the aforementioned dual task.

Our main thesis of the present paper is that a DM who adopts the proposed counterfactual strategy IC_{I_S} does not exhibit AA, thus escaping Ellsberg's paradox. The main results of this paper may now be stated as follows: a) For every coherent (additive) counterfactual prior probability function P_0^c, the posterior probability function P_1, which is generated from P_0^c via Bayesian conditionalization on I_S, is coherent. b) If DM's probability function P_1 is coherent, and her utility function is increasing, then it is impossible for DM to exhibit Ellsberg-type behaviour.

3 Ellsberg's Paradox

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 DM is given the specific information I_S : "30 balls are red and the remaining 60 balls are either black or yellow in unknown proportion". DM will draw a ball at random, which means that each ball has an equal probability of being drawn.

DM 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:

Outcomes

	red is drawn	black is drawn	yellow is drawn
f	100	0	0
g	0	100	0
f^*	100	0	100
g^*	0	100	100

In the framework of SEUM, 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 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 DM believes that drawing a red ball is more probable than drawing a black ball, then DM's probabilistic sophistication requires her to believe that drawing a red or yellow ball is more probable than drawing a black or yellow ball. Therefore, SEUM theory entails the following conditional proposition: "if 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 SEUM theory. Moreover, such a pair of choices imply that DM's subjective probability function is not additive which means that DM is not probabilistically sophisticated.

Our main aim in the present section is to demonstrate the following: If DM adopts the counterfactual strategy IC_{I_S} described in Introduction, then her current probabilistic beliefs, that is, those that have taken into account the specific information I_S, are coherent. This in turn implies that whether DM turns out to be probabilistically sophisticated or not, depends entirely on whether she decides to process I_S in the *indirect* way outlined in the previous section or not.

We assume that DM's epistemic life begins at $t = 1$. At this time DM receives the specific information I_S : "30 of 90 balls in the urn are red". This means that DM's interest in the phenomenon under study, coincides with the time at which she came to know I_S. This information provides DM with the objective probabilities of some (but not all) the elements of the relevant space $\mathcal{F}$, thus inducing the following partition $(\mathcal{F}_k, \mathcal{F}'_k)$ of the algebra of propositions $\mathcal{F}$:

$$\mathcal{F}_k = \{s_R, s_{BY}, s_{RBY}, \perp\},$$

$$\mathcal{F}'_k = \{s_B, s_Y, s_{RB}, s_{RY}\},$$

where

$$\begin{aligned} s_R &: \text{"a red ball is drawn"} \\ s_B &: \text{"a black ball is drawn"} \\ s_Y &: \text{"a yellow ball is drawn"} \\ s_{RB} &: \text{"a red or a black ball is drawn"} \\ s_{RY} &: \text{"a red or a yellow ball is drawn"} \\ s_{BY} &: \text{"a black or a yellow ball is drawn"} \\ s_{RBY} &: \text{"a red or a yellow or a black ball is drawn"} \\ \perp &: \text{"no ball is drawn".} \end{aligned} \tag{5}$$

The probabilistic content of I_S takes the form of the following objective probabilities,

$$Ch(s_R) = \frac{1}{3},$$

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

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

Remark 1 *It is important to note that the information I_S refers to the objective probabilities of the elements of $\mathcal{F}_k$. Whether DM endorses these objective probabilities as her own subjective probabilities is another question. Most philosophers agree that rationality dictates that DM should conform to a "probability coordination principle", according to which DM adopts as her own subjective probabilities the corresponding objective ones, provided that the latter are known (see, for example, Strevens 2017). The most important of such principles is David Lewi's "Principal Principle" (PP, Lewis 1976). In the analysis that follows we tacitly assume that the DM adheres to PP.*

Given that the DM receives the specific information I_S at time $t = 1$, which coincides with the beginning of DM's epistemic life, suggests that DM does not possess an actual prior probability function P_0^a, formed at some previous time, (say) $t = 0$. In the absence of P_0^a, DM is assumed to make the counterfactual move and "mentally go back" to $t = 0$, in which I_S was not certain, but instead it was one of the many alternative pieces of information (information propositions) that DM could receive at $t = 1$. At that hypothetical moment, DM deliberates her probabilistic assignments on $\mathcal{F}$ relativized only with respect to the background information, I_B, available at that moment. In other words, DM asks herself the question "what would my prior probability function P_0^c be, had I only known I_B?". By asking herself this question, DM is effectively searching her Carnapian prior, or as Meacham more recently called it, "Ur-Prior". This is defined as "the credences a subject should have if she had no evidence, a subject's initial credences, a subject's evidential standards, and any function that plays the right diachronic role." (2016, pp. 1). The important thing to note is that at that hypothetical moment, DM is "uniformly ignorant" about the objective probabilities of $\mathcal{F}$. Hence, in contemplating P_0^c, DM does not enter the cognitive state of comparative ignorance, which as already mentioned, is considered to be the main cause of ambiguity aversion.

Before we proceed any further, the following clarification is in order: When DM is instructed "to ignore I_S", she is advised to do so only temporarily. No canon of rationality would tell DM to discard useful information. Indeed, failure to take I_S into account means that DM is doomed to "commit the most obvious inconsistency of reasoning." (Jaynes 1968, pp. 227). What we emphatically suggest is that the proper effect of I_S on DM's probabilistic beliefs at $t = 1$, be elicited via Bayesian conditionalization.

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 : This is 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 (being the content of I_S), 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".

Let us now assume that DM has moved back to time $t = 0$, in which she counterfactually possesses only I_B. This means that at that moment, DM has decided to (temporarily) ignore the specific information I_S : "30 of 90 balls in the urn are red". Under this epistemic state, DM is judging her subjective probability function P_0^c that is about to assign on $\mathcal{F}$. Part of the background information I_B, with respect to which P_0^c is relativized, is the proposition that "the number of all possible color combinations that may be accommodated in the urn is 4,186". Specifically, based on I_B alone, DM knows that the urn may contain 0 red, 0 black and 90 yellow balls, or 0 red, 1 black and 89 yellow balls, etc. In other words, at $t = 0$ the agent knows that one of the following "theoretical propositions" is true:

$$\begin{aligned}
&\mathcal{H}_{(0,0,90)} : \text{"The urn contains 0 red, 0 black and 90 yellow balls"} \\
&\mathcal{H}_{(0,1,89)} : \text{"The urn contains 0 red, 1 black and 89 yellow balls"} \\
&\mathcal{H}_{(0,2,88)} : \text{"The urn contains 0 red, 2 black and 88 yellow balls"} \\
&\quad . \\
&\quad . \\
&\quad . \\
&\mathcal{H}_{(0,90,0)} : \text{"The urn contains 0 red, 90 black and 0 yellow balls"} \\
&\quad . \\
&\quad . \\
&\quad . \\
&\mathcal{H}_{(90,0,0)} : \text{"The urn contains 90 red, 0 black and 0 yellow balls"}.
\end{aligned} \tag{6}$$

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 $t = 0$, DM does not know which proposition of $\mathbf{H}$ is the true one. As a result, she treats the elements of $\mathbf{H}$ as part of the domain of P_0^c. This in turn implies that 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 (5) 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 $\mathcal{F}_{ext}^0$ is a Boolean algebra. It must be noted that although DM is interested in the propositions of $\mathcal{F}^0_{ext}$, the space that contains the propositions of "betting interest" for DM remains $\mathcal{F}$, $\mathcal{F} \subset \mathcal{F}^0_{ext}$.

Let us now examine the question of whether DM's counterfactual epistemic condition, assumed at $t = 0$, is conducive to DM's forming a coherent counterfactual subjective probability function P^c_0, defined on $\mathcal{F}^0_{ext}$. The important thing to note is that at time $t = 0$, DM is in a state of uniform rather than comparative ignorance. In other words, at time $t = 0$ she is equally uninformed about the elements of $\mathcal{F}^0_{ext}$. This means that all the elements of $\mathcal{F}^0_{ext}$ are equally "vague" or equally "ambiguous", meaning that DM's epistemic reliability is distributed uniformly over them. But if there is no "comparative ambiguity" in DM's mind, concerning the elements of $\mathcal{F}^0_{ext}$, there are no grounds for assuming that DM is compelled to exhibit ambiguity aversion. Equivalently, there is no basis for assuming that DM will prove to be probabilistically non-sophisticated. Of course, even in this counterfactual environment, DM might still be unable to express precise probabilities for each and every element of $\mathcal{F}^0_{ext}$. Uniform ambiguity can also be harmful as far as DM's ability to judge her own probabilities is concerned. However, this is a different source of probabilistic deficiency, which falls outside the scope and spirit of Ellsberg's paradox.[2] Indeed, in the context of the Ellsberg paradox, DM's behaviour is studied under the assumption that DM knows the asymmetric information I_S all along, that is, she is always under the state of comparative ignorance. If this assumption is withdrawn, and DM is assumed to have no probabilistic information whatsoever, then DM is transferred into the state of uniform ignorance. In this state it is not unrealistic to claim that DM would adopt the non-informative prior P^c_0, according to which each color has equal probability of being drawn. This means that Ellsberg's paradox may be thought of as a case in which DM's probabilistic sophistication is easier to achieve under uniform ignorance rather than comparative ignorance.

Since P^c_0 obeys the rules of probability calculus, it also satisfies the law of total probability, according to which $\forall a \in \{R, B, Y\}$,

$$P^c_0(s_a) = \sum_{\mathbf{i} \in \boldsymbol{I}} P^c_0(s_a \mid \mathcal{H}_{\mathbf{i}}) P_0(\mathcal{H}_{\mathbf{i}}). \tag{7}$$

To calculate $P^c_0(s_a \mid \mathcal{H}_{\mathbf{i}})$, i.e. the probability of drawing a $a-$colored ball conditional on the hypothesis $\mathcal{H}_{\mathbf{i}}$, it is convenient to define $\boldsymbol{I}^k_a \subset \boldsymbol{I}$, 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, $P^c_0(s_a \mid \mathcal{H}_{\mathbf{i}})$ is non-zero if $\mathbf{i} \in \boldsymbol{I}^k_a$. Moreover, $P^c_0(s_a \mid \mathcal{H}_{\mathbf{i}}) = \frac{k}{90}$, if $\mathbf{i} \in \boldsymbol{I}^k_a$ and $card(\boldsymbol{I}^k_a) = 91 - k$, $\forall k = 0, ..., 90$ and

[2] The question of whether DM's probabilistic beliefs are or have to be rational (in the sense that they obey the rules of mathematical probability) has always been at the heart of the (still ongoing) debate between Bayesianists and their critics.

$\forall a \in \{R, B, Y\}$. As a result,

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

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 $t = 0$ 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. $I_{\mathbf{s}}$ denotes the specific information that DM may receive in the future. As already mentioned, this information 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, $I_{\mathbf{s}} = I_S = \boldsymbol{I}_R^{30}$, with I_S being the information that "30 of the 90 balls are red".[3] 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 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 $P_0^c(s_a \mid I_S)$. This is because in the context of her counterfactual reasoning, the factual proposition I_S must be treated on a par with any other possible, (but hypothetical) information scenario $I_{\mathbf{s}} \in \mathbf{H} - \{I_S\}$.

Let us now calculate DM's conditional probabilities $P_0^c(s_a \mid I_S)$.[4] Using (**??**) we have that $\forall a \in \{R, B, Y\}$:

$$P_0^c(s_a \mid I_S) = \sum_{\mathbf{i} \in \boldsymbol{I}} P_0^c(s_a \mid \mathcal{H}_{\mathbf{i}} \wedge I_S) P_0^c(\mathcal{H}_{\mathbf{i}} \mid I_S), \tag{8}$$

where

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

and

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

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

[4]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.

Therefore,

$$P_0^c\left(s_a \mid I_S\right) = \sum_{\mathbf{i}\in \boldsymbol{I}} P_0^c\left(s_a \mid \mathcal{H}_{\mathbf{i}} \wedge I_S\right) P_0^c\left(\mathcal{H}_{\mathbf{i}} \mid I_S\right) = \sum_{\mathbf{i}\in \boldsymbol{I}_R^{30}} P_0^c(s_a \mid \mathcal{H}_{\mathbf{i}})\frac{1}{61}, \forall a \in \{R, B, Y\}. \tag{9}$$

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

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

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, (9) takes the form,

$$P_0^c\left(s_a \mid I_S\right) = \sum_{k=0}^{60} \sum_{\mathbf{i}\in \boldsymbol{I}_R^{30} \cap \boldsymbol{I}_a^k} P_0^c(s_a \mid \mathcal{H}_{\mathbf{i}})\frac{1}{61} = \sum_{k=0}^{60} \frac{k}{90}\frac{1}{61} = \frac{1}{3}, a \in \{B, Y\}. \tag{10}$$

It follows that,

$$P_0^c\left(s_R \mid I_S\right) = P_0^c\left(s_B \mid I_S\right) = P_0^c\left(s_Y \mid I_S\right) = \frac{1}{3} \tag{11}$$

and

$$P_0^c\left(s_{RY} \mid I_S\right) = P_0^c\left(s_{BY} \mid I_S\right) = P_0^c\left(s_{RB} \mid I_S\right) = \frac{2}{3}. \tag{12}$$

At this point, DM concludes the calculations for her subjective probabilities of the propositions of betting interest, namely those in $\mathcal{F}$, conditional on I_S. As already mentioned, DM is supposed to repeat the procedure outlined above in order to calculate her subjective conditional probabilities for the remaining information scenarios, $I_{\mathbf{s}} \in \mathbf{H} - \{I_S\}$. Once this is done, DM's formation of her own counterfactual prior P_0^c is completed.

Now it is time for DM to exit the counterfactual mode of probabilistic thinking and mentally return to the actual time point $t = 1$. In other words, she is ready to implement the second step of IC_{I_S}. This step is particularly simple. All that DM has to do is merely to adopt the counterfactual conditional probabilities, given in (11) and (12), as her current subjective probabilities for $t = 1$. By doing so, DM ends up with the following additive subjective probability function P_1, defined on $\mathcal{F}$:

$$P_1(s_R) = P_1(s_B) = P_1(s_Y) = \frac{1}{3} \text{ and } P_1(s_{RY}) = P_1(s_{BY}) = P_1(s_{RB}) = \frac{2}{3}.$$

The detailed analysis of the counterfactual strategy IC_{I_S}, presented above, might have made clear the two-step character of this strategy. The first step is the "information processing" one, whereas the second is the step of "commitment". Indeed, all that I_S has to say about the propositions in $\mathcal{F}$ is analyzed by

DM in the first step, and is captured by her conditional probabilities $P_0^c(s_a \mid I_S)$, $a \in \{R, B, Y\}$. In the second step, there is no information processing activity involved. Instead, all that DM is asked to do is to abide by her subjective conditional probabilities $P_0^c(s_a \mid I_S)$, $a \in \{R, B, Y\}$ calculated in the first step, thus adopting them as her current probabilities $P_1(s_a)$ for $t = 1$.

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 I_S the actual information were $I_{S'}$: "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 $P_1(s_R) = P_0^c(s_R \mid I_S)$ 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. This generalized result is stated in the form of the following proposition:

Proposition 2 *(**??**) 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_1 of time $t = 1$, generated by $P_1(A) = P_0^c(A \mid I_{\mathbf{s}})$, $A \in \mathcal{F}$ is coherent (i.e. additive).*

Proof. See Appendix. ■

The above Proposition demonstrates the following: If DM follows the two-step, indirect, counterfactual way of processing any specific information that may come to know $t = 1$ then her subjective probability function at $t = 1$ is additive. This means that DM is probabilistically sophisticated. As such, she does not exhibit Ellsberg-type choices. To this end, we prove the following proposition:

Proposition 3 *(**??**) If DM at $t = 1$ is probabilistically sophisticated (that is, for any coherent subjective probability function P_1) and her utility function is increasing, then it is impossible for her to exhibit Ellsberg-type choices.*

Proof. See Appendix. ■

The two propositions, presented above, comprise a potential solution to Ellsberg's paradox. The question which arises at this point is what is the normative and the descriptive status of the counterfactual strategy IC_{I_S} upon which the resolution of the paradox is based. The normative appeal of IC_{I_S} is supported by the arguments presented in Sections 2 and 3. Additional reasons for adopting IC_{I_S} will be furnished in Sections 5 and 6. As far as the descriptive validity of IC_{I_S} is concerned, the question is whether normal (rather than ideal) people would be willing to implement it, after they were presented with its normative virtues. Answering this question requires the design and implementation of a relevant empirical study, in which subjects will be asked whether, under the proper guidance, find IC_{I_S} appealing.

4 Arguments in Support of Counterfactual Priors

Even if one accepts the main result of the present paper, namely that IC_{I_S} eliminates DM's ambiguity, (and hence Ellsberg's paradox) the question remains: why is it normatively appealing to adopt IC_{I_S}? Is IC_{I_S} exhausted in solving Ellsberg's paradox alone, or are there more fundamental reasons which dictate the adoption of such a strategy? Below we identify psychological, methodological and logical reasons for the adoption of counterfactual probabilistic reasoning.

4.1 Psychological Arguments for IC_{I_S}

The first psychological argument was offered, surprisingly, not by a psychologist but a philosopher, namely Rudolph Carnap, admittedly one of the greatest philosophers of science of the twentieth century. Carnap (1962, 1971) argues that DM's pure or genuine probabilistic beliefs are expressed by what he calls "permanent dispositions to believe", which are identified with DM's initial credence function Cr_0. The latter is defined as the singular subjective probability function that DM has (or would have) prior to the acquisition of any evidence (that is at the beginning of DM's epistemic life). This initial credence function stands in sharp contrast with the current probabilistic beliefs, CPB_n, that DM happens to have at some point in time n. CPB_n do not capture the true probabilistic dispositions of DM, but instead they codify DM's "momentary inclinations to believe" at time n. How do DM's current probabilistic beliefs, Cr_n, at time n inherit the "trait of DM's underlying permanent intellectual character"? This can only be achieved if DM conditionalizes on all the information accumulated between $t = 0$ and $t = n$, using Cr_0 as the relevant vehicle (see Carnap 1971, pp. 18-19).[5] The main message from Carnap's suggestion is the following: If 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 I_B. If DM does so, then her current beliefs, Cr_n 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 $I_B \cup I_S$, then these beliefs (CPB_n) are temperamental or capricious and almost surely different than Cr_n.

Apart from Carnap, there are many other philosophers who have suggested DM's endorsement of an initial subjective probability function, that has to be formulated without the direct influence of any specific information I_S, even if DM actually knows I_S, (see Lewis 1980, Levi 1980, Skyrms 1983, and more recently Meacham 2008 and Titelbaum 2013). Howson (1991) in particular, strongly recommends the deletion of any specific information I_S from the body

[5] Carnap's Cr_0, CPB_n and Cr_n correspond to P_0^c, $P_1^{I_S}$, and P_1, respectively as defined in previous sections of the present paper..

of information, upon which DM's prior probability is based, as an answer to "the problem of old evidence". The latter was introduced by Glymour (1980) and Gardner (1982) and was initially interpreted as a serious pitfall of Bayesian Confirmation Theory (BCT). In line with Carnap, Howson argues that in order for BCT to get off the ground, "the dispositional properties of the agent's belief structure", as reflected in P_0^c, must be identified (Howson 1984, pp. 246).

At the heart of the Carnapian argument lies the view that in order for DM to identify her true probabilistic dispositions, she must bring herself in a psychological state in which I_S is not treated as certainty (even if DM actually knows I_S), but rather as one of the many alternative, yet unrealized, possibilities. Put differently, the appropriate psychological state, which DM should pursue, in order to evaluate her prior probability function is the one in which none of the aforementioned possibilities has been realized. This, of course, raises the question of why such a state is the appropriate one. An answer to this question may be given by the following example: Suppose that DM, being at $t = 1$, contemplates her probability of the event A : "I will live for another five years". At that time, DM learns the information I_S : "I am just diagnosed with lung cancer". To this end, DM has two options: (a) DM attempts to evaluate her subjective probability of A, under the psychological burden provoked by her viewing I_S as certain. In this case, she comes up with $P_1^{I_S}(A) = p_1$. (b) DM evaluates her probability of A counterfactually by asking herself the question "what would my probability of A be, were I to know that I_S is true?" In this case, DM treats I_S 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. DM's probability of A in this environment is represented by $P_0^c(A \mid I_S) = p_2$. It seems reasonable to assume that $p_1 > p_2$.

Another example of how DM's actual encounter with I_S 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).

A third example of this kind comes from Greek mythodology.[6] Ulysses knows already from $t = 0$ that when he will listen to sirens' song at $t = 1$, 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 $t = 1$ 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 DM's overall ability

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

to think objectively. Indeed, there is a plethora of empirical studies that document a negative relationship between DM's level of anxiety (which in our case is caused by DM's perception of I_S 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 I_S may be thought of as a "situational moderator", which negatively affects DM's information processing skills (see Humphreys and Revelle 1984). A common implication of both interpretations is the following: if DM treats I_S as certain (that is when $P_1(I_S) = 1$), then she may experience emotional biases, which in turn impair her ability to uncover her genuine probabilistic dispositions.

Additional evidence, dictating the necessity of temporarily removing I_S from the corpus of (certain) knowledge with respect to which DM's prior is relativized, is offered by studies that examine whether DM alters her behavior when she is presented with convincing counterarguments highlighting her "mistakes" (see MacCrimmon and Larsson 1979, Slovic and Tversky 1974). These studies find that DM continues to prefer risk over uncertainty (that is continues to exhibit AA) even when the associated inconsistencies are revealed to her. This in turn implies that as long as DM treats I_S as "certain", she does not think of her AA preferences as a "mistake" that should be corrected under the proper guidance. Instead, she appears to be defending her initial choices, which means that if DM is irrational, she is persistently (or stubbornly) so.[7]

4.2 Methodological Arguments for IC_{I_S}

Before we develop our methodological arguments in favor of IC_{I_S}, let us begin with those related to the less controversial case in which DM processes I_S by means of a pre-existing actual prior probability function P_0^a. Hereafter, this strategy will be referred to as IA_{I_S}. In this setup, DM's epistemic life covers two periods, namely $t = 0$ and $t = 1$. As already mentioned, the specific information I_S becomes available at $t = 1$. This implies that of the two aforementioned periods, only the second one, $t = 1$, is characterized by information asymmetry between $\mathcal{F}_k$ and $\mathcal{F}'_k$. Hence, only at $t = 1$, the conditions are conducive for DM to experience ambiguity and develop ambiguity aversion. On the contrary, at $t = 0$ DM is equally ignorant about the probabilities of $\mathcal{F}_k$ and $\mathcal{F}'_k$. Due to this informational symmetry at $t = 0$, DM is not prone to entering the state of ambiguity. In the absence of ambiguity, we may assume (in a standard Bayesian fashion) that at $t = 0$, DM possesses an actual prior probability function P_0^a. This in turn implies that the subjective conditional probabilities $P_0^a(A \mid I_S)$, $A \in \mathcal{F}$ are all determined at $t = 0$. Moving to $t = 1$, DM faces two options: Either she forms her new probabilistic beliefs via Bayesian conditionalization, $P_1(A) = P_0^a(A \mid I_S)$, $A \in \mathcal{F}$, or forgoes her previous probabilistic commitments, thus embodying the newly acquired I_S into her current subjective probabilities $P_1^{I_S}(A)$, $A \in \mathcal{F}$ in a discretionary and probably nebulous manner. The

[7] Of course, the question of whether DM successfully implements IC_{I_S} when she is instructed to do so is an empirical one. What we stress in this paper is the normative force of the proposed counterfactual strategy.

advantages of the first over the second option are well known: Bayesian conditionalization ensures that (a) DM remains probabilistically sophisticated at $t = 1$, and thus not vulnerable to a static Dutch book, (b) DM is dynamically consistent and thus not vulnerable to a diachronic Dutch book. Hence, in the presence of P_0^a, the only way for DM to exhibit ambiguity aversion (at $t = 1$) is to abstain from her probabilistic commitments made at $t = 0$, that is to be dynamically inconsistent.[8]

The best strategy for the identification of the methodological merits of IC_{I_S} is to compare it with IA_{I_S}.The question to be answered is whether, even in the presence of an actual P_0^a, there is still scope for DM to employ a counterfactual prior P_0^c as the suitable vehicle for her probabilistic updating.

This question may seem bizare at first sight. Why should DM eschew her own readily available P_0^a and reconstruct counterfactually another prior probability function? Besides, the best cognitive state for DM to form her prior probability is the state in which I_S is actually rather than counterfactually contingent. Admittedly, this is the major advantage of IA_{I_S} over IC_{I_S}. This advantage notwithstanding, there are cases in which IC_{I_S} may serve DM's purpose of updating her probabilistic beliefs better than IA_{I_S}. More specifically, the potential preference of IC_{I_S} over IA_{I_S} may be grounded on the following arguments:

(i) Error Correction

Assume that at $t = 0$, DM constructs her actual prior P_0^a. This means that part of her probabilistic assessments at this point are the conditional probabilities $P_0^a(A \mid I_S)$, $A \in \mathcal{F}$. Assume further that DM is a committed Bayesian, which means that at $t = 1$, when she learns that I_S is true, she updates her beliefs according to the Bayesian rule, $P_1(A) = P_0^a(A \mid I_S)$, $A \in \mathcal{F}$. Now assume that at $t = 1$, DM comes to believe that some of her conditional probabilities, $P_0^a(A \mid I_S)$ formed at $t = 0$ are mistaken. What should she do? Bayesian conditionalization is a rule that implies DM's perpetual commitment to her initial probabilistic assessments, thus allowing for no error correction at any time in the future. Prior conditional probabilities are rigid, which means that the probabilistic impact of I_S on A is determined once and for all at $t = 0$. Regardless of whether DM changes her mind, at $t = 1$, about some of her prior conditional probabilities, strict adherence to the Bayesian rule compels her to update according to $P_1(A) = P_0^a(A \mid I_S)$, thus allowing her initial error to contaminate her current probabilistic beliefs. In their criticism of conditionalization, Bacchus, Kyburg and Thalos (1990) remark: "We cannot assign coherent probabilities to all statements at the drop of a hat; it requires some reflection, some computation, and even then *we must be prepared to have made a mistake. We must be able to back up and reconsider.*" (1990, pp 484, emphasis added).

At the heart of the aforementioned criticism lies the elapsed time between

[8] In the case of IC_{I_S}, there is only one epistemic time point for DM, namely $t = 1$. As a result, the issue of dynamic consistency does not arise in the first place: there are no DM's probabilistic commitments at $t = 0$ (since there is no epistemic $t = 0$) to be maintained or violated at $t = 1$. This means that IC_{I_S} cannot be defended on the grounds of standard diachronic Dutch book arguments, such as the ones available for IA_{I_S}.

the period $t = 0$ at which DM's prior probability is formed and the period $t = 1$ at which the decision is made. DM's cognitive state at $t = 0$ may be different than that at $t = 1$, with this difference being uncorrelated with the arrival of I_S. Put differently, at $t = 1$, DM may find herself compelled to state that $P_1(A) \neq P_0^a(A \mid I_S)$ not because of the arrival of I_S, but rather because between $t = 0$ and $t = 1$ she had the opportunity to further reflect on her prior beliefs and decide that these beliefs are represented not by $P_0^a(A \mid I_S)$ but rather by $P_{00}^a(A \mid I_S)$, with $P_0^a(A \mid I_S) \neq P_{00}^a(A \mid I_S)$ for some $A \in \mathcal{F}$. In such a case, DM does not abandon Bayesian conditionalization altogether, but rather she changes her mind about her prior probability function. When does this change of opinion occur? Obviously, at $t = 1$. But then the switch from P_0^a to P_{00}^a is tantamount to the switch from P_0^a to P_0^c. In other words, P_{00}^a can be nothing else but DM's counterfactual prior P_0^c. Hence, we have identified a reason for which P_0^c is superior to P_0^a, namely the property that P_0^c as opposed to P_0^a allows for "error correction".

The preceding discussion gives rise to the following question: Is a rational DM allowed to change her mind? More specifically, what is DM's rationality status if at $t = 1$ she relinquishes P_0^a in favor of P_{00}^a (or, equivalently P_0^c)? Is it possible that this dynamic inconsistency between ex-ante planning and ex-post decision has some adverse pragmatic consequences for DM? In particular, is such a DM vulnerable to a dynamic Dutch book? The answer to this question is negative. In order for a cunning bookie to be able to construct a diachronic Dutch book against DM, he must know in advance, i.e. from $t = 0$, DM's probabilities at $t = 1$. This of course implies that DM herself must know at $t = 0$ her future probabilities for $t = 1$, and, in addition, to be willing to announce these probabilities at $t = 0$. Obviously, this is not the case when DM changes her probabilistic beliefs between $t = 0$ and $t = 1$. In other words, at $t = 0$ DM believes that her future probabilities will be $P_1(A) = P_0^a(A \mid I_S)$ but at $t = 1$, these initial beliefs are falsified ex post and her current probabilities turn out to be $P_1(A) = P_{00}^a(A \mid I_S)$.

Of course, such a change of opinion raises questions about DM's rationality which are grounded on another aspect of her probabilistic behaviour, namely her inability to accomplish an "ex ante probabilistic plan" and execute it. To this end, Lewis (1997) remarks: "If you can't tell in advance how your beliefs would be modified by a certain course of experience, that also is a kind - a different kind - of irrationality on your part". (1997, pp 407).

(ii) Unconceived Alternatives

Another case in which IC_{I_S} displays a vital advantage over IA_{I_S} is that of "unconceived alternatives". The Problem of Unconceived Alternatives (PUA), which is considered to be a serious challenge to Bayesian epistemology, was initially put forward by Duhem (1954), and revived more recently by Stanford (2006). PUA may be thought of as a special case of a more general epistemological problem, namely the underdetermination of scientific theories by empirical data. As such, it contributed to a pre-existing skepticism about the "realism of scientific theories", namely that our best current theories are (at least) approximately true.

Stanford introduces PUA as follows: "I propose what I will call *the new induction over the history of science*: that we have, throughout the history of scientific inquiry and in virtually every scientific field, repeatedly occupied an epistemic position in which we could conceive of only one or a few theories that were well confirmed by the available evidence, while subsequent inquiry would routinely (if not invariably) reveal further, radically distinct alternatives as well confirmed by the previously available evidence as those we were inclined to accept on the strength of that evidence." (2006, pp. 19). Stanford presents a number of distinct examples for the history of Physics and Biology, which exhibit the pattern of PUA: "These prominent examples at least suggest a robust, distinctive pattern, in which the available evidence cited in support of each earlier theory ultimately turned out to support one or more competitors *unimagined at the time* just as well." (2006, pp. 20, emphasis added).

How does PUA affect the comparison between IC_{I_S} and IA_{I_S}? As will be shown below, under PUA, IC_{I_S} is not only a better option than IA_{I_S}, it is the only option. Let us again consider two time periods, $t = 0$ and $t = 1$. Assume that at $t = 0$, DM possesses an actual prior probability function, P_0^a, defined on $\mathcal{F}_{ext}^0$. The space $\mathcal{F}_{ext}^0$ contains all the theoretical propositions H_i, $i = 1, 2, ..., m$ that DM has conceived as possible explanations of her empirical data as well as all the relevant empirical propositions (alternative courses of experience), E_j, $j = 1, 2, ..., k$ that may come to be true in the future. However, when $t = 1$ comes, DM's realizes that the space of epistemic possibilities is larger than she originally perceived at $t = 0$. In particular, at this time point, DM becomes aware of an additional theoretical proposition, H_{m+1}, or/and a new empirical proposition E_{k+1}, which were ignored (by DM) at $t = 0$.[9] This means that the "unconceived alternatives" H_{m+1} and/or E_{k+1} do not belong to the original propositional space $\mathcal{F}_{ext}^0$, which implies that these propositions (together with all the compound propositions in which they participate) do not carry prior actual probabilities $P_0^a(H_{m+1})$ and/or $P_0^a(E_{k+1})$, respectively. As a result, at $t = 1$ DM feels compelled to replace $\mathcal{F}_{ext}^0$ with the extended space $\mathcal{F}_{ext}^1$, $\mathcal{F}_{ext}^0 \subset \mathcal{F}_{ext}^1$ and define a new prior P_0^c on $\mathcal{F}_{ext}^1$. Since the specific information I_S becomes known at $t = 1$, it follows that DM's new prior has to be a counterfactual one. It is important to note that under the present scenario of PUA, DM cannot, at $t = 1$, employ her old actual prior P_0^a as her vehicle of conditionalization, even if she wanted to do so. The emergence of the new alternatives (unimagined at $t = 0$) results in $P_0^a(A) \neq P_0^c(A)$ even for propositions that belong to the original space $\mathcal{F}_{ext}^0$. This means that DM finds at $t = 1$ her original probabilities of the "conceived propositions", namely the elements of $\mathcal{F}_{ext}^0$ to be entirely different. For example, assume that at $t = 1$, DM realizes that the Ellsberg urn contains apart from red, black and yellow balls, (say) green balls as well, with the total number of balls in the urn now being equal to 100. As a result, DM realizes that she has to expand her original space of possibilities from $\mathcal{F}_{ext}^0$ to $\mathcal{F}_{ext}^1$, with the difference $\mathcal{F}_{ext}^1 - \mathcal{F}_{ext}^0$ containing

[9] Rowbottom (2019) answers the question of why the empirical proposition E_{k+1} might be unconceived as follows: "The observations in question are theory-laden, and the necessary theory (or set of theories) to conceive of them is itself unconceived." (2019, pp. 3950)

all the "green-related" propositions. This expansion, however, does not leave the original probabilities defined on $\mathcal{F}_{ext}^{0}$ unaffected. For example, the original probability of drawing a red ball conditional on I_S is equal to 1/3, whereas the corresponding probability with respect to $\mathcal{F}_{ext}^{1}$ is 3/10. This means that PUA makes the original P_0^a entirely obsolete. Hence, the construction of a new P_0^c is the only viable option for DM, if she wishes to accommodate the newly conceived alternative.

PUA-type instances in economic decision making is abundant. The adoption of quantitative easing programs by the major central banks, following the Great Financial Crisis of 2008-09 and the introduction of a negative deposit facility interest rate by the European Central Bank in 2014, are cases that were not conceived as possible little time before their occurrence.

5 The Logical Structure of Counterfactual Probabilistic Reasoning

It must have been clear by now that the main normative recommendation of the present paper is that any specific information I_S must be processed counterfactually, as outlined by the IC_{I_S} strategy. In this section, we analyze the logical structure of the concept of "information processing" in general, showing that counterfactual reasoning is in fact the main element of this structure.

Let us assume that DM wishes to analyze the effect of the event I on the event A, that is to deduce the implications of the information I for the event of interest A. These implications depend on the "connection" between the events I and A. Hence, DM is primarily interested in identifying the connection (causal, statistical or logical) between the event I (the information bearer) and the event A that she is ultimately interested in. In trying to identify this connection, DM asks herself the question: "What does it mean for A *if* I occurs?". Alternatively, stated in propositional language, the previous question may take the form: "*If* the proposition I is true, then is the proposition A true as well?". This means that the connection between I and A takes the form of the conditional "if I then A" or "$I \longrightarrow A$", where "$\longrightarrow$" is a conditional connective (for more on the exact nature of this connective, see below). In this context, DM has to decide about her own subjective probability $P(I \longrightarrow A)$ of the conditional $I \longrightarrow A$. Assume that DM knows that there is a universal law stating that whenever an event of type I occurs, an event of type A follows. In this case, DM is entitled to set $P(I \longrightarrow A)$ equal to unity. In such a case, the connection between I and A is deterministic or the strength of the connection is complete. Alternatively, assume that no such universal law exists. Instead, there is a statistical (or probabilistic) law, stating that whenever an event of type I occurs, an event of type A *usually* follows. Moreover, assume that the statistical law is precise enough to enable quantification of the term "usually". For example, "whenever an event of type I occurs an event of type A follows in 70 percent of the cases" (frequency interpretation) or "whenever an event of type I occurs an event of type A has propensity equal to 0.7 to follow" (propensity interpretation). In this case, DM guided by the Principal Principle, defers to the relevant objective probability, thus setting $P(I \longrightarrow A)$ equal to 0.7. Again, DM is entitled to believe that there is some connection between I and A, which although not deterministic, is probabilistically strong.

Let us analyze in more detail the structure of the conditional proposition $I \longrightarrow A$. As already mentioned, the interpretation of $I \longrightarrow A$ in the natural language takes the form of "if...then". Within the compound proposition I is not known to be true, even if $I \longrightarrow A$ were known to be true. Indeed, in order for DM to deduce the truth of A, apart from the truth of $I \longrightarrow A$, she is also required to know the truth of I. That is, she needs another proposition stating that I is true. Then, $I \longrightarrow A$ and I can be conjoined to form the premises of a valid argument (*modus ponens*) whose conclusion is the proposition A. This means that within the context of $I \longrightarrow A$, I is not treated as certainty but rather as contingency, which in turn implies that the investigation of the connection between I and A structurally involves counterfactual or hypothetical thinking. In other words, DM's counterfactual mode of reasoning is mandatory rather than optional, dictated by the logical structure of the problem at hand. As Evans et. al. (2007) remarks: "'*If*' is used to initiate the imagination and simulation of possibilities, a process that we term *hypothetical thinking"* (2007, pp. 1772). In a similar vein, Edgington (1986) emphasizes the hypothetical status of I in the analysis of $I \longrightarrow A$: "It is necessary to suppose (or assume) that some epistemic possibility is true, and to consider what else would be the case, or *would be likely to be the case, given this supposition.* The conditional expresses the outcome of such thought processes." (1986, pp. 4, emphasis added).

The preceding analysis implies that DM's problem of identifying the "strength of the connection" between I and A is reduced to the assessment of her subjective probability $P(I \longrightarrow A)$ of the conditional $I \longrightarrow A$. This in turn raises the following question: How should DM determine $P(I \longrightarrow A)$? To this end, DM may feel tempted to equate this probability with the conditional probability $P(A \mid I)$ of the consequent A given the antecedent I. Such a disposition seems to be supported by the so-called Ramsey's test. Ramsey (1929) suggests the following rule of evaluating an indicative conditional: "If two people are arguing 'If p will q?' and are both in doubt as to p, they are adding p *hypothetically* to their stock of knowledge and arguing on that basis about q." (1929, pp. 155, emphasis added). Further support for equating $P(I \longrightarrow A)$ with $P(A \mid I)$ is offered by many empirical psychological studies, suggesting that this is in fact the manner in which most people evaluate the probability of a conditional proposition. Evans et. al. (2007) summarizes these findings as follows (where their propositions p and q correspond to our I and A, respectively): "They (the subjects under examination) do this by hypothetically supposing p and then running a mental simulation in which they evaluate q" (2007, pp. 1772).

However, despite its descriptive validity, the question whether, from a logical point of view, this is the correct way of evaluating the probability of the conditional $I \longrightarrow A$ remains. To put it succinctly, does the equality

$$P(I \longrightarrow A) = P(A \mid I),\ P(I) > 0 \tag{13}$$

always hold? That is, does (13) hold for any coherent P and any propositions I and A in the domain of P? This question has been a central theme in the "philosophy of the conditionals" literature since the early 1960s, causing a lot of debate up to the present time. The current consensus seems to suggest that although (13) does not always hold (in fact, it holds only in some "trivial" cases), a variation of (13) is the most satisfactory account of the logic of conditionals. This variation, usually referred to as Adams' Thesis (Adams 1965, 1970, 1990), was offered by Adams as a reaction to Lewis's "triviality results" (see Lewis 1976). At the heart of Adams' Thesis lies the assumption that the conditional connective "$\longrightarrow$" is not truth-functional. This means that $I \longrightarrow A$ is not a truth-functional proposition and as such it cannot be "probability bearer". In other words, $P(I \longrightarrow A)$ is meaningless just because $I \longrightarrow A$ is not a proposition. This however, generates the following question. If $P(I \longrightarrow A)$ is meaningless, then what happens to the left-hand side of (13)? Adams' answer to this question is the following: Although we cannot talk about the probability of $I \longrightarrow A$, that is we cannot talk about the probability that the proposition $I \longrightarrow A$ is true, nevertheless we can meaningfully speak about the *assertability* $As(I \longrightarrow A)$ of the conditional $I \longrightarrow A$.[10] How may $As(I \longrightarrow A)$ be interpreted? It is definitely not probability, but it is not entirely alien to it. Although $As(\cdot)$ is not $P(\cdot)$, which immediately means that Lewis's assumption (L-i) is relaxed, it nevertheless retains some of the features of $P(\cdot)$. Before, we discuss the properties of $As(\cdot)$, let us formally state Adams' Thesis:

$$As(I \longrightarrow A) = P(A \mid I),\ P(I) > 0. \tag{14}$$

(14) reads as "the assertability of the indicative conditional $I \longrightarrow A$ always goes by the conditional probability of the consequent given the antecedent". By depriving $As(\cdot)$ of the formal properties of $P(\cdot)$, Adams makes (14) immune to Lewis's triviality results.

However, so far we have said what $As(\cdot)$ is not, namely that it is not a proper probability. Now it is time to say what $As(\cdot)$ actually is or at least what $As(\cdot)$ looks like. Unfortunately, Adams has not given a clear answer to this question. Hajek (2011) informs us (based on his private communication with Adams) that what Adams had in mind when he introduced the term was "reasonableness of belief" rather than "appropriateness of utterance". This means that although $I \longrightarrow A$ lacks truth values, it nevertheless has assertabilities (or degrees of assertability) which, though not formal probabilities, are something similar to them. It must be emphasized that the right hand-side of (14) continues to represent formal probabilities, albeit conditional ones. This in turn implies that DM's system of beliefs for non-conditional propositions is still represented

[10] Jackson (1987) suggested that "assertibility" (with an "i") is a better term than "assertability" (with an "a") and Hayek (2012) proposed the term "acceptability".

by a coherent subjective probability function. When it comes to conditionals, however, DM invokes a different representation of beliefs, namely assertabilities rather than probabilities. Yet, the values of these assertabilities are determined - according to Adams' Thesis - by the corresponding conditional probabilities. As a result, the known properties of the term in the right-hand side of (14) may illuminate us about the unknown properties of the term in the left-hand side of this equation (for example, we may deduce that $As(I \longrightarrow A) \in [0, 1]$).

What are the implications of the preceding analysis for the normative status of the counterfactual strategy IC_{I_S} proposed in this paper? The most important element of the foregoing discussion is the following: Even if the specific information I_S is a certain event (that is an event that has already occurred or a proposition that is known to be true) at the time of DM's decision making (in our framework at $t = 1$), DM has to counterfactually treat I_S as a contingency rather than an actuality. This is because, DM is not interested in the event I_S per se, but rather in the implications of I_S for the event(s) of interest A. In order to judge these implications, DM has to determine the probability or the assertability of the conditional $I_S \longrightarrow A$. By its very logical structure, the information proposition I_S (the antecedent) in the conditional $I_S \longrightarrow A$ is "only supposed not accepted" (see Arlo-Costa 1999, Raidl 2019). Hence, DM has to decide about the asertability of $I_S \longrightarrow A$ on the supposition that the truth value of I_S is not yet determined. In the context of Adams' Thesis, this means that the status of I_S in the left-hand side of (14) is hypothetical. An immediate implication of this is that I_S is also hypothetical in the right-hand side of (14). Hence, the right-hand side of (14) represents a counterfactual conditional probability.

6 Conclusions

In this paper, we proposed a novel strategy for dealing with ambiguity aversion (AA) and the resulting Ellsberg-type choices. We first identified the presence of "asymmetric information" as the main cause of ambiguity aversion. In particular, we have emphasized the fact that in order for DM to exhibit AA, a specific piece of probabilistic information I_S, must be present, which in turn induces a partition $\{\mathcal{F}_k, \mathcal{F}'_k\}$ of the relevant space $\mathcal{F}$ in DM's mind. Specifically, I_S informs DM about the objective probabilities of $\mathcal{F}_k$ only. Hence, DM enters the cognitive state of comparative ignorance, in which she feels more competent to bet on events/propositions of $\mathcal{F}_k$ than on those of $\mathcal{F}'_k$. This analysis has established the following causal chain: The asymmetric information I_S causes DM to feel ignorant of the events in $\mathcal{F}'_k$ compared to those in $\mathcal{F}_k$, which in turn triggers the feeling of ambiguity, thus causing DM to exhibit ambiguity aversion.

Based on the identification of the aforementioned causal chain, we focused on the modal status of I_S and especially on DM's epistemic attitude towards it. There, we put forward the idea that the way in which I_S is allowed to affect DM's probabilistic beliefs depends on whether DM thinks of I_S as certainty or contingency. To this end, we argued that in order for DM to identify her true "dispositions to believe", that is her genuine credence function, she must treat I_S 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 DM's own standards) a non zero probability of being true. This mode of information processing was coined "indirect counterfactual" way of processing I_S, (symbolized as IC_{I_S}) and constitutes the main proposal of this paper. IC_{I_S} is a two-step procedure: In the first step, DM judges her prior probability function P_0^c without the direct influence of I_S, whereas in the second step she utilizes I_S by conditionalizing on it, using P_0^c as the required vehicle. This means that IC_{I_S} combines "the best of both worlds": it allows DM to exploit the information content of I_S, while at the same time, prevents the process of information extraction to contaminate her true probabilistic dispositions.

Once comparative ignorance and the associated ambiguity aversion have been removed, the road towards solving Ellsberg's paradox is open. In fact, the solution of Ellsberg's paradox, produced in this paper, emerges as a direct application of the proposed counterfactual strategy IC_{I_S} to the specific choice problem described by Ellsberg. The fact that such a simple solution of Ellsberg's paradox has eluded the large literature on ambiguity aversion is, in and of itself, another paradox.

A great deal of the paper has been devoted in analyzing the psychological, methodological and logical merits of IC_{I_S}. Is IC_{I_S} a procedure that is "tailor made" for solving Ellsberg's paradox? Or are there any reasons, over and above its efficacy in solving Ellsberg's paradox, that make IC_{I_S} normatively appealing? The answers to these questions offer an unequivocal support for the implementation of IC_{I_S} not only in the narrow case of Ellsberg's paradox, but in any case in which DM attempts to ascertain her probabilistic beliefs in the presence of both background and specific information. The psychological arguments in favor of IC_{I_S} have already been spelled out: As Carnap realized more than seventy years ago, processing I_S counterfactually enables DM to identify her true probabilistic dispositions rather than her whimsical, momentary inclinations. Treating I_S as one of the many alternative information scenarios that may come true "in the future", DM brings herself in a neutral psychological state in which her probabilistic judgments are not affected by either the positive or the negative emotions that the inevitability of I_S is likely to generate.

As far as the methodological advantages of IC_{I_S} are concerned, we showed that IC_{I_S} may be employed to solve two of the main complaints about Bayesian Confirmation Theory, namely the problems of "error correction" and "unconceived alternatives". These advantages were unearthed by comparing IC_{I_S} with IA_{I_S}, with the latter being referred to the probabilistic processing of I_S in the case that DM already possesses an actual prior at the time that she learns the truth of I_S. We argued that IA_{I_S} does not work in the case that DM realizes that her actual prior probabilistic assessments were wrong and/or, that an alternative theoretical or empirical proposition was not conceived at the time that she was forming her actual prior. In such case, DM is advised to answer the following question: "What would have been my prior, had I known the error or had I conceived the alternative?" This type of counterfactual thinking may drive DM out of her predicament and provide her with a new "corrected" prior, albeit a counterfactual one.

Finally, we put forward a logical argument for IC_{I_S} and especially for the counterfactual status of the conditional probability that looms in IC_{I_S}. This argument is based on the analysis of the logical structure of the concept of "information processing". The latter amounts to deriving the implications of the proposition I_S for the proposition of interest A. These implications are captured by the connection between I_S and A, with the latter being summarized by the conditional $I_S \rightarrow A$. The strength of this connection (as evaluated by DM) is quantified by the subjective probability $P(I_S \rightarrow A)$, or in a less strict form by the assertability $As(I_S \rightarrow A)$. Assuming the soundness of Adams' Thesis, the assertability of $I_S \rightarrow A$ is equal to the probability of A conditional on I_S. Given that I_S in $I_S \rightarrow A$ is definitely in hypothetical mood, the corresponding conditional probability $P(A \mid I_S)$, whose nature is determined by the modal status of I_S, can be no other than a counterfactual one.

References

Adams, E. (1965). The logic of conditionals. *Inquiry*, 8, 166–197.

Adams, E. (1975). The logic of conditionals. Dordrecht: Reidel.

Adams, E. (1990). Review of Conditionals, by Frank Jackson. *Philosophical Review*, 99(3), 433–435.

Arlo-Costa, H. (1999). Belief revision conditionals: basic iterated systems. *Annals of Pure and Applied Logic*, 96(1-3), 3-28.

Bacchus, F., Kyburg, H. E., Thalos, M. (1990). Against conditionalization. *Synthese*, 475-506.

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.

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.

Duhem, P. (1954). The Aim and Structure of Physical Theory, translated from 2nd ed. by P. W. Wiener; originally published in 1914 as La Theorie Physique: Son Objet et sa Structure (Paris: Marcel Riviera & Cie.), Princeton, NJ: Princeton University Press.

Edgington, D. (1986). Do conditionals have truth conditions?. *Critica: Revista Hispanoamericana de Filosofia*, 3-39.

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.

Evans, J. S. B., Handley, S. J., Neilens, H., and Over, D. E. (2007). Thinking about conditionals: A study of individual differences. *Memory & cognition*, 35(7), 1772-1784.

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

Gardner, M. R. (1982). Predicting novel facts. *The British Journal for the Philosophy of Science*, 33(1), 1-15.

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

Glymour, C. (1980). Theory and Evidence. Princeton: Princeton University Press.

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

Hajek, A. (2011). Conditional probability. In P. Bandyopadhyay & M. Forster (Eds.), *Handbook for philosophy of statistics*, Amsterdam: Elsevier

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.

Howson, C. (1984). Bayesianism and Support by Novel Facts. *The British Journal for the Philosophy of Science,* 35, 245-251.

Howson, C. (1991). The 'Old Evidence Problem', *The British Journal for the Philosophy of Science,* 42(4), 547–55.

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.

Jackson, F. (1980). Conditionals and possibilia. *Proceedings of the Aristotelian Society*, 81, 125-137.

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

391-418.

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.

Levi, I. (1980), The Enterprise of Knowledge. Cambridge: The MIT Press.

Levi, I. (1985). Imprecision and indeterminacy in probability judgment. *Philosophy of Science,* 52, 390-409.

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

Lewis, D. (1994). Humean Supervenience Debugged. *Mind*, 103(412), 473–490.

Lewis, D. (1980). A Subjectivist's Guide to Objective Chance, in Richard C. Jeffrey (ed.) *Studies in Inductive Logic and Probability*, Vol II., 263–294.

MacCrimmon, K. R. and Larson, S. (1979). Utility theory: axioms versus paradoxes. *Expected Utility Hypotheses and the Allais Paradox* (ed. Allais, M. and Hagen, O.), 333–410.

Meacham, C. J. (2008). Sleeping Beauty and the Dynamics of De Se Beliefs. *Philosophical Studies,* 138, 245–70.

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

Raidl, E. (2019). Completeness for counter-doxa conditionals–using ranking semantics. *The Review of Symbolic Logic*, 12(4), 861-891.

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

Rowbottom, D.P. (2019). Extending the argument from unconceived alternatives: observations, models, predictions, explanations, methods, instruments, experiments, and values. *Synthese,* 196, 3947–3959.

Ramsey, F. P. (1929). General Propositions and Causality. *Philosophical Papers*, edited by D. H. Mellor, 1990, 145–63.

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

Schmeidler, 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.

Slovic, P. and Tversky, A. (1974). Who accepts Savage's axiom?. *Behavioral science*, 19(6), 368-373.

Stanford, P. K. (2006). Exceeding Our Grasp: Science, History, and the Problem of Unconceived Alternatives, New York: Oxford University Press.

Titelbaum, M. G. (2013). Quitting certainties: A Bayesian framework modeling degrees of belief. Oxford University Press.

7 Appendix

Proof of Proposition (??):

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". 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 $\forall a \in \{R, B, Y\}$:

$$P_1(s_a) = P_0^c(s_a \mid I_S) = \sum_{\mathbf{i} \in \boldsymbol{I}} P_0^c(s_a \mid \mathcal{H}_{\mathbf{i}} \wedge I_S) P_0^c(\mathcal{H}_{\mathbf{i}} \mid I_S), \tag{15}$$

where

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

and

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

Therefore,

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

Clearly,

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

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

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

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,

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

Therefore,

$$P_1(s_a) = \sum_{k=0}^{90-l} \sum_{\mathbf{i}\in \boldsymbol{I}^l_R \cap \boldsymbol{I}^k_a} P_0^c(s_a \mid \mathcal{H}_{\mathbf{i}}) P_0^c\left(\mathcal{H}_{\mathbf{i}} \mid I_S\right) =$$

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

As a result,

$$P_1(s_R) = \frac{l}{90},$$
$$P_1(s_B) = \sum_{k=0}^{90-l} \frac{k}{90} \sum_{\mathbf{i}\in \boldsymbol{I}_R^l \cap \boldsymbol{I}_B^k} \frac{p_{\mathbf{i}}}{\sum_{\mathbf{j}\in \boldsymbol{I}_R^l} p_{\mathbf{j}}} = \frac{E_0(s_B \mid I_S)}{90} \text{ and}$$
$$P_1(s_{RY}) = \frac{l}{90} + \sum_{k=0}^{90-l} \frac{k}{90} \sum_{\mathbf{i}\in \boldsymbol{I}_R^l \cap \boldsymbol{I}_Y^k} \frac{p_{\mathbf{i}}}{\sum_{\mathbf{j}\in \boldsymbol{I}_R^l} p_{\mathbf{j}}} = \frac{l + E_0(s_Y \mid I_S)}{90},$$
$$P_1(s_{BY}) = \frac{(90-l)}{90}.$$

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

Proof of Proposition (??):

The expected utility of f and g is given by:

$$YAX(f) = E(U(f)) = P_1(s_R) \times U(100) + P_1(s_{BY}) \times U(0)$$
$$YAX(g) = E(U(g)) = P_1(s_{RY}) \times U(0) + P_1(s_B) \times U(100).$$

Similarly, the expected utility of f^* and g^* is given by:

$$YAX(f^*) = E(U(f^*)) = P_1(s_{RY}) \times U(100) + P_1(s_B) \times U(0)$$
$$YAX(g^*) = E(U(g^*)) = P_1(s_R) \times U(0) + P_1(s_{BY}) \times U(100)$$

We have already proved that, under Laplace's "principle of indifference",

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

Therefore, for any increasing utility function U, an agent will choose f over g and f^* over g^* iff

$$l > 30.$$

Similarly, she will choose g over f and g^* over f^* iff

$$l < 30.$$

Finally, she will be indifferent between f and g and between f^* and g^* iff

$$l = 30.$$

Hence, there is no case under which a paradox emerges.

Similarly, for the general case where

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

where $p_{\mathbf{i}} > 0$ and $\sum_{\mathbf{i}\in\boldsymbol{I}} p_{\mathbf{i}} = 1$, we have proved that

$$P_1(s_R) = \frac{l}{90},$$
$$P_1(s_B) = \sum_{k=0}^{90-l} \frac{k}{90} \sum_{\mathbf{i}\in\boldsymbol{I}_R^l \cap \boldsymbol{I}_B^k} \frac{p_{\mathbf{i}}}{\sum_{\mathbf{j}\in\boldsymbol{I}_R^l} p_{\mathbf{j}}} = \frac{E_0(s_B \mid I_S)}{90} \text{ and}$$
$$P_1(s_{RY}) = \frac{l}{90} + \sum_{k=0}^{90-l} \frac{k}{90} \sum_{\mathbf{i}\in\boldsymbol{I}_R^l \cap \boldsymbol{I}_Y^k} \frac{p_{\mathbf{i}}}{\sum_{\mathbf{j}\in\boldsymbol{I}_R^l} p_{\mathbf{j}}} = \frac{l + E_0(s_Y \mid I_S)}{90},$$
$$P_1(s_{BY}) = \frac{90-l}{90}.$$

Note that $P_1(s_{RY}) = \frac{l+E_0(s_Y|I_S)}{90} = \frac{90-E_0(s_B|I_S)}{90}$. Again, for any increasing utility function U, an agent chooses f to g and f^* to g^* iff

$$l > E_0(s_B \mid I_S).$$

Similarly, she will choose g to f and g^* to f^* iff

$$l < E_0(s_B \mid I_S).$$

Finally, she will be indifferent between f and g and between f^* and g^* iff

$$l = E_0(s_B \mid I_S).$$

Hence, there is no case under which a paradox emerges, even if the priors are not formed under the "principle of indifference".