Can Bayesian agents always be rational? A principled analysis of consistency of an Abstract Principal Principle*

Zalán Gyenis†
Alfréd Rényi Institute of Mathematics
13-15 Reáltanoda str. 1053 Budapest, Hungary
gyz@renyi.hu

Miklós Rédei ‡
Department of Philosophy, Logic and Scientific Method
London School of Economics and Political Science
Houghton Street, London WC2A 2AE, UK
m.redei@lse.ac.uk

November 5, 2013

Abstract

The paper takes the Abstract Principal Principle to be a norm demanding that subjective degrees of belief of a Bayesian agent be equal to the objective probabilities once the agent has conditionalized his subjective degrees of beliefs on the values of the objective probabilities, where the objective probabilities can be not only chances but any other quantities determined objectively. Weak and strong consistency of the Abstract Principal Principle are defined in terms of classical probability measure spaces. It is proved that the Abstract Principal Principle is weakly consistent and that it is strongly consistent in the category of probability measure spaces where the Boolean algebra representing the objective random events is finite. It is argued that it is desirable to strengthen the Abstract Principal Principle by adding a stability requirement to it. Weak and strong consistency of the resulting Stable Abstract Principal Principle are defined, and the strong consistency of the Abstract Principal Principle is interpreted as necessary for a non-omniscient Bayesian agent to be able to have rational degrees of belief in all epistemic situations. It is shown that the Stable Abstract Principal Principle is weakly consistent, but the strong consistency of the Stable Abstract Principal principle remains an open question. We conclude that we do not yet have proof that Bayesian agents can have rational degrees of belief in every epistemic situation.

1 The claims

The aim of this paper is to investigate the consistency of what we call the "Abstract Principal Principle". We take the Abstract Principal Principle to be a general norm that regulates probabilities representing the subjective degrees of beliefs of an abstract Bayesian agent by requiring the agent's degrees of beliefs to be equal to the objective probabilities if the agent knows the values of the objective probabilities. We call this principle the *Abstract* Principal Principle because nothing is assumed about the specific nature of the objective probabilities — they can be (finite or infinite) relative frequencies, chances, propensities, ratios of some sort, or any other quantities viewed as determined objectively, i.e. independently of the agent and his beliefs.

After stating the Abstract Principal Principle informally in section 2, we describe in a non-technical way the consistency problem to be analyzed in the paper. The consistency in question is

*This paper was presented in the LSE Choice Group Workshop "Recent work on the Principal Principle", LSE, October 18, 2013. We thank the audience for their comments.
†Research supported in part by the Hungarian Scientific Research Found (OTKA). Contract number: K83726.
‡Research supported in part by the Hungarian Scientific Research Found (OTKA). Contract number: K100715.

of fundamental nature: it expresses the harmony of the Abstract Principal Principle with the basic structure of measure theoretic probability theory. It will be seen that this consistency comes in different degrees of strength and we develop them step by step, proceeding from weaker to stronger. In section 3 we define formally, in terms of classical measure theoretic probability theory specified by the standard Kolmogorovian axioms the *weak consistency* of the Abstract Principal Principle (Definition 1) and prove that the Abstract Principal Principle is weakly consistent (Proposition 1). The proof will reveal a weakness in the concept of weak consistency and in section 4 we will define the *strong consistency* of the Abstract Principal Principle. We then prove (details of the proof are given in the Appendix) that, under some (mild) assumptions on the agent's prior subjective probability, the Abstract Principal Principle is strongly consistent in the category of probability spaces where the Boolean algebra representing the random events having objective probability has a finite number of elements (Proposition 2). We will then argue that it is very natural to strengthen the Abstract Principal Principle by requiring it to satisfy a *stability* property, which expresses that conditional degrees of beliefs in events *already* equal (in the spirit of the Abstract Principal Principle) to the objective probabilities of the events do not change as a result of conditionalizing them further on knowing the objective probabilities of *other* events (in particular of events that are independent with respect to their objective probabilities). We call this amended principle *Stable* Abstract Principal Principle (if stability is required only with respect to further conditionalizing on values of probabilities of *independent* events: *Independence-Stable* Principal Principle). This stability requirement leads to suitably modified versions of both the weak and strong consistency of the (*Independence*-)Stable Abstract Principal Principle (Definitions 4 and 5). We will prove that the Stable Abstract Principal Principle is weakly consistent (Proposition 3), irrespective of cardinality of the Boolean algebra of random events. This entails that the Independence-Stable Abstract Principal Principle also is weakly consistent (Proposition 4). (Details of the proof are given in the Appendix.)

The *strong* consistency of both the Stable and of the Independence-Stable Abstract Principal Principle remain open problems however, even in the category of probability spaces with a finite Boolean algebra. In section 7 we will interpret the strong consistency of the Stable Abstract Principal Principle as a necessary condition for a non-omniscient Bayesian agent to be able to have rational degrees of belief under all epistemic conditions and thus will conclude that we do not yet have proof that Bayesian agents can in principle always be rational.

Throughout the systematic part of the paper containing the results (sections 2-7) no references will be given to relevant and related literature. Section 8 puts the results into context. In particular, we discuss in this section the relevance of the notion of strong consistency of the Stable Abstract Principal Principle from the perspective of the Principal Principle about chances. The main message of this section is that the strong consistency of the Stable Abstract Principal Principle also is necessary for the consistency of both the original formulation of the Principal Principle by Lewis and for the consistency of some of the subsequent modifications of the original Principal Principle that have been proposed in the literature on the Principal Principle. Since the consistency of the Stable Abstract Principal Principle is an open problem, it is not known at this point whether the original Principal Principle and of some of the suggested modifications are in harmony with the basic conceptual structure of measure theoretic probability.

2 The Abstract Principal Principle informally

The Abstract Principal Principle regulates probabilities representing the subjective degrees of beliefs of an abstract Bayesian agent by stipulating that the subjective degrees of beliefs $p_{subj}(A)$ of the agent in events A are related to the objective probabilities $p_{obj}(A)$ as

$$p_{subj}(A|\ulcorner p_{obj}(A) = r\urcorner) = p_{obj}(A) \tag{1}$$

where $\ulcorner p_{obj}(A) = r\urcorner$ denotes the proposition "the objective probability, $p_{obj}(A)$, of A is equal to r".

The Abstract Principal Principle – and the formulation given by eq. (1) in particular – presupposes that both p_{subj} and p_{obj} have the features of a probability measure: they both are assumed to be additive maps defined on a Boolean algebra taking values in the unit interval $[0, 1]$: p_{obj} is supposed to be defined on a Boolean algebra $\mathcal{S}_{obj}$ of random events viewed as specified objectively (equivalently, on a Boolean algebra of propositions stating that the random events happen); and p_{subj} also is supposed to be a map with a domain of definition being a Boolean algebra $\mathcal{S}_{subj}$.

It is crucial to realize that the Boolean algebras $\mathcal{S}_{obj}$ and $\mathcal{S}_{subj}$ cannot be unrelated: for the conditional probability $p_{subj}(A|\ulcorner p_{obj}(A) = r\urcorner)$ in eq. (1) to be well-defined via the Bayes'rule, it is necessary that the Boolean algebra $\mathcal{S}_{subj}$ serving as the domain of definition of the probability measure p_{subj} contains *both* the Boolean algebra $\mathcal{S}_{obj}$ of random events *and* with every random event A also the proposition $\ulcorner p_{obj}(A) = r\urcorner$ — for if $\mathcal{S}_{subj}$ does not contain both of these two sorts of propositions then the formula $p_{subj}(A|\ulcorner p_{obj}(A)=r\urcorner)$ cannot be interpreted as an expression of conditional probability specified by the Bayes'rule because the conditional probability $p_{subj}(A|\ulcorner p_{obj}(A)=r\urcorner)$ given by the Bayes'rule reads

$$p_{subj}(A|\ulcorner p_{obj}(A)=r\urcorner) = \frac{p_{subj}(A\cap\ulcorner p_{obj}(A)=r\urcorner)}{p_{subj}(\ulcorner p_{obj}(A)=r\urcorner)} \tag{2}$$

thus all three propositions A, $\ulcorner p_{obj}(A)=r\urcorner$ and $A\cap\ulcorner p_{obj}(A)=r\urcorner$ must belong to the Boolean algebra $\mathcal{S}_{subj}$ on which the subjective probability p_{subj} is defined.

It is far from obvious however that, given *any* Boolean algebra $\mathcal{S}_{obj}$ of random events with *any* probability measure p_{obj} on $\mathcal{S}_{obj}$, there exists a Boolean algebra $\mathcal{S}_{subj}$ meeting these algebraic requirements in such a way that a probability measure p_{subj} satisfying the condition (2) also exists on $\mathcal{S}_{subj}$. If there exists a Boolean algebra $\mathcal{S}^*_{obj}$ of random events with a probability measure p^*_{obj} giving the objective probabilities of events for which there exists *no* Boolean algebra $\mathcal{S}_{subj}$ on which a probability function p_{subj} satisfying (2) can be defined, then the Abstract Principal Principle would not be maintainable in general – the Abstract Principal Principle would then be inconsistent as a general norm: In this case the agent, being in the epistemic situation of facing the objective facts represented by $(\mathcal{S}^*_{obj}, p^*_{obj})$, cannot have degrees of belief satisfying the Abstract Principal Principle for fundamental structural reasons inherent in the basic structure of classical probability theory. We say that the Abstract Principal Principle is *weakly consistent* if it is *not* inconsistent in the sense described. (The adjective "weakly" will be explained shortly.) To formulate the weak consistency of the Abstract Principal Principle precisely, we fix some notation and recall some definitions first.

3 Weak consistency of the Abstract Principal Principle

The triplet $(X,\mathcal{S},p)$ denotes a classical probability measure space specified by the Kolmogorovian axioms, where X is the set of elementary random events, $\mathcal{S}$ is a Boolean algebra of (some) subsets of X representing a general event and p is a probability measure on $\mathcal{S}$. Given two Boolean algebras $\mathcal{S}$ and $\mathcal{S}'$, the map $h\colon\mathcal{S}\to\mathcal{S}'$ is a Boolean algebra embedding if it preserves all Boolean operations:

$$h(A_1\cup A_2) = h(A_1)\cup h(A_2) \tag{3}$$

$$h(A_1\cap A_2) = h(A_1)\cap h(A_2) \tag{4}$$

$$h(A^\perp) = h(A)^\perp \tag{5}$$

and is injective:

$$A\neq B \text{ entails } h(A)\neq h(B) \tag{6}$$

If $h\colon\mathcal{S}\to\mathcal{S}'$ is a Boolean algebra embedding, then $h(\mathcal{S})$ is an isomorphic copy of $\mathcal{S}$, which can be viewed as a Boolean subalgebra of $\mathcal{S}'$. From the perspective of probability theory elements A and $h(A)$ can be regarded as identical $A\leftrightarrow h(A)$.

The probability space $(X',\mathcal{S}',p')$ is called an extension of the probability space $(X,\mathcal{S},p)$ with respect to h if h is a Boolean algebra embedding of $\mathcal{S}$ into $\mathcal{S}'$ that preserves the probability measure p:

$$p'(h(A)) = p(A) \qquad A\in\mathcal{S} \tag{7}$$

Definition 1. The Abstract Principal Principle is defined to be *weakly consistent* if the following hold: Given any probability space $(X_{obj},\mathcal{S}_{obj},p_{obj})$, there exists a probability space $(X_{subj},\mathcal{S}_{subj},p_{subj})$ and a Boolean algebra embedding h of $\mathcal{S}_{obj}$ into $\mathcal{S}_{subj}$ such that

(i) For every $A\in\mathcal{S}_{obj}$ there exists an $A'\in\mathcal{S}_{subj}$ with the property

$$p_{subj}(h(A)|A') = p_{obj}(A) \tag{8}$$

(ii) If $A,B\in\mathcal{S}_{obj}$ and $A\neq B$ then $A'\neq B'$.

The intuitive content of Definition 1 should now be clear: The probability space $(X_{obj},\mathcal{S}_{obj},p_{obj})$ describes the objective probabilities; in particular $p_{obj}(A)$ is the probability of the random event $A\in\mathcal{S}_{obj}$. The Boolean algebra $\mathcal{S}_{subj}$ in the probability space $(X_{subj},\mathcal{S}_{subj},p_{subj})$ contains not only the "copies" $h(A)$ of all the random events $A\in\mathcal{S}_{obj}$ (together with all the undistorted algebraic relations among the random events), but, with every random event $A\in\mathcal{S}_{obj}$, also an element A' to be interpreted as representing the proposition "the objective probability, $p_{obj}(A)$, of A is equal to r" (this proposition we denoted by $\ulcorner p_{obj}(A)=r\urcorner$). If $A\neq B$ then $A'\neq B'$ must be the case because $\ulcorner p_{obj}(A)=r\urcorner$ and $\ulcorner p_{obj}(B)=s\urcorner$ are different propositions – this is expressed by (ii) in the definition. The main content of the Abstract Principal Principle is then expressed by condition (8), which states that the *conditional* degrees of beliefs $p_{subj}(h(A)|A')$ of an agent about random events $h(A)\leftrightarrow A\in\mathcal{S}_{obj}$ are equal to the objective probabilities $p_{obj}(A)$ of the random events, where the condition A' is that the agent knows the values of the objective probabilities.

Proposition 1. *The Abstract Principal Principle is weakly consistent.*

The above proposition follows from the proposition that states the weak consistency of the *Stable* Abstract Principal Principle (Proposition 3), which we state later and prove in the Appendix. To motivate the need to strengthen the notion of weak consistency, we give here a proof of Proposition 1 under the simplifying restriction that the Boolean algebra $\mathcal{S}_{obj}$ has a finite number of elements. The proof will expose the conceptual weakness of the notion of weak consistency; recognizing the weakness leads naturally to the notion of strong consistency.

Let $(X_{obj}\mathcal{S}_{obj}, p_{obj})$ be a probability space and assume that $\mathcal{S}_{obj}$ is a finite Boolean algebra having $n < \infty$ elements. Let X_n be any set having n elements and $f\colon \mathcal{S}_{obj} \to X_n$ be a bijection. Let p be any probability measure on the power set $\mathcal{P}(X_n)$ of X_n such that $p(A) \neq 0$ for any $\emptyset \neq A \in \mathcal{P}(X_n)$. Consider the standard product probability space

$$(X_{obj} \times X_n, \mathcal{S}_{obj} \otimes \mathcal{P}(X_n), p_{obj} \times p) \tag{9}$$

of the probability spaces $(X_{obj}, \mathcal{S}_{obj}, p_{obj})$ and $(X_n, \mathcal{P}(X_n), p)$, where $p_{obj} \times p$ is the product measure on $\mathcal{S}_{obj} \otimes \mathcal{P}(X_n)$. Recall that the Boolean algebra $\mathcal{S}_{obj} \otimes \mathcal{P}(X_n)$ is the smallest Boolean algebra on $X_{obj} \times X_n$ that contains all the sets of the form $A \times B$ with $A \in \mathcal{S}_{obj}$, $B \in \mathcal{P}(X_n)$; and for $A \times B$, with $A \in \mathcal{S}_{obj}$ and $B \in \mathcal{P}(X_n)$ we have

$$(p_{obj} \times p)(A \times B) = p_{obj}(A)p(B) \tag{10}$$

The maps h and g defined by

$$\mathcal{S}_{obj} \ni A \mapsto h(A) \doteq (A \times X_n) \in \mathcal{S}_{obj} \otimes \mathcal{P}(X_n) \tag{11}$$

$$\mathcal{P}(X_n) \ni B \mapsto g(B) \doteq (X_{obj} \times B) \in \mathcal{S}_{obj} \otimes \mathcal{P}(X_n) \tag{12}$$

are Boolean algebra embeddings and $p_{subj} \doteq (p_{obj} \times p)$ is a probability function that satisfies eq. (8) with

$$A' = X_{obj} \times \{f(A)\} \tag{13}$$

because

$$p_{subj}(h(A)|A') = \frac{p_{subj}((A \times X_n) \cap (X_{obj} \times \{f(A)\})}{p_{subj}(X_{obj} \times \{f(A)\})} = \frac{(p_{obj} \times p)(A \times \{f(A)\})}{(p_{obj} \times p)(X_{obj} \times \{f(A)\})} \tag{14}$$

$$= \frac{p_{obj}(A)p(\{f(A)\})}{p_{obj}(X_{obj})p(\{f(A)\})} = p_{obj}(A) \tag{15}$$

In short, given $(X_{obj}, \mathcal{S}_{obj}, p_{obj})$, the product probability space (9) satisfies the conditions required of $(X_{subj}, \mathcal{S}_{subj}, p_{subj})$ by the definition of weak consistency of the Abstract Principal Principle, and the reason for this is that the restriction of the probability measure $p_{subj} = (p_{obj} \times p)$ to the (isomorphic copy $h(\mathcal{S}_{obj})$ of) Boolean algebra $\mathcal{S}_{obj}$ of random events *coincides* with the probability measure p_{obj} giving the objective probability of the random events, and, as the calculation (14) shows, conditionalizing with respect to the propositions $A' = X_{obj} \times \{f(A)\}$ (which are interpreted as stating the values of objective probabilities) does not change $p_{subj}(A)$ because the propositions A' and $A \leftrightarrow h(A) = (A \times X_n)$, in virtue of their lying in different components of the product Boolean algebra, are independent with respect to the product measure $p_{subj} = (p_{obj} \times p)$ on the product algebra. That is to say, in the situation represented by this product state case, the agent's degrees of beliefs $p_{subj}(A)$ are equal to the objective probabilities $p_{obj}(A)$ *without* any conditionalizing, and conditionalizing them on the independent propositions stating the values of the objective probabilities does not change the already correct degrees of belief. Clearly, this is a very exceptional situation however, and it is more realistic to assume that the agent's degrees of belief are only equal to the objective probabilities *after* conditionalizing them on knowing the values of objective probabilities but *not* before. One would like to know if the Abstract Principal Principle is possible to maintain (is consistent) under these more realistic circumstances. The definition of weak consistency does not say anything about this more stringent consistency however; thus it is desirable to strengthen it in the manner specified in the next section.

4 Strong consistency of the Abstract Principal Principle

Definition 2. The Abstract Principal Principle is defined to be *strongly consistent* if the following hold: Given any probability space $(X_{obj}, \mathcal{S}_{obj}, p_{obj})$ and another probability measure p^0_{subj} on $\mathcal{S}_{obj}$, there exists a probability space $(X_{subj}, \mathcal{S}_{subj}, p_{subj})$ and a Boolean algebra embedding h of $\mathcal{S}_{obj}$ into $\mathcal{S}_{subj}$ such that

(i) For every $A \in \mathcal{S}_{obj}$ there exists an $A' \in \mathcal{S}_{subj}$ with the property

$$p_{subj}(h(A)|A') = p_{obj}(A) \tag{16}$$

(ii) If $A, B \in \mathcal{S}_{obj}$ and $A \neq B$ then $A' \neq B'$.

(iii) The probability space $(X_{subj}, \mathcal{S}_{subj}, p_{subj})$ is an extension of the probability space $(X_{obj}, \mathcal{S}_{obj}, p^0_{subj})$ with respect to h; i.e. we have

$$p_{subj}(h(A)) = p^0_{subj}(A) \qquad A \in \mathcal{S}_{obj} \tag{17}$$

The above Definition 2 differs from the definition of weak consistency (Definition 1) only by the additional requirement (iii). The intuitive content of this requirement is that the agent's prior probability function p_{subj} restricted to the random events can be equal to an arbitrarily chosen measure p^0_{subj} on $\mathcal{S}_{obj}$; in particular the agent's prior subjective probabilities about random events can differ from the objective probabilities of the random events given by p_{obj}.

Definition 3. An agent's prior degrees of beliefs represented by p^0_{subj} are called *non-extreme* (with respect to the objective probability function p_{obj}) if they satisfy the two conditions below:

- The agent's prior probabilities are *not* zero in events that have *non-zero* objective probability:
$p^0_{subj}(A) = 0$ entails $p_{obj}(A) = 0$
- The agent's prior probabilities are *not* equal to 1 in events whose non-occurrence have a *nonzero* objective probability:
$p^0_{subj}(A) = 1$ entails $p_{obj}(A) = 1$

Proposition 2. *The Abstract Principal Principle is strongly consistent in the category of probability spaces with a finite Boolean algebra under the further assumption that the agent's prior degrees of beliefs are not extreme.*

We prove the above proposition in the Appendix. The proof is based on a technique that constructs, in finite steps, an extension of a probability space with a finite Boolean algebra; the extension itself will be a probability space in which the Boolean algebra has a finite number of elements. The extension procedure differs from taking the standard product of the probability measure space to be extended with a suitable chosen other one, which was the technique used to prove the weak consistency of the Abstract Principal Principle (Proposition 1), and which also is the technique we will use to prove the weak consistency of the Stable Abstract Principal Principle later. We conjecture that the strong consistency cannot be proved by taking the standard product as an extension; furthermore we conjecture that Proposition 2 holds generally:

Conjecture 1. *Proposition 2 is true* without *the assumption that the Boolean algebra of random events is finite.*[1]

5 Strengthening the Abstract Principal Principle: The Stable Abstract Principal Principle

Once the agent has learned the values of the objective probabilities and has adjusted his subjective degrees of belief by conditionalizing on this evidence, $p_{subj}(h(A)|\ulcorner p_{obj}(A) = r\urcorner) = r$, he can in principle conditionalize *again* his already conditionalized degrees of belief; he can in particular conditionalize on the evidence of values of objective probabilities of events different from A: on $\ulcorner p_{obj}(B) = s\urcorner$, say. What should be the result of this second conditionalization? Since the agent's conditional degrees of belief $p_{subj}(h(A)|\ulcorner p_{obj}(A) = r\urcorner)$ in A are already correct, i.e. equal to the objective probabilities, learning an additional *truth*, namely the value of the objective probability $p_{obj}(B)$, it would be an irrational move on the agents's part to change his already correct degree of belief about A. That is to say, a *rational* agent's conditional subjective degrees of belief should be *stable* in the sense of satisfying the following condition:

$$p_{subj}\big(h(A)|\ulcorner p_{obj}(A) = r\urcorner\big) = p_{subj}\big(h(A)|\ulcorner p_{obj}(A) = r\urcorner \cap \ulcorner p_{obj}(B) = s\urcorner\big) \quad (\forall B \in \mathcal{S}_{obj}) \tag{18}$$

Another reason why stability should be a feature of the conditional subjective degrees of belief is the following. If A and B are independent with respect to their objective probabilities $p_{obj}(A \cap B) =$

[1]See the paper [4].

$p_{obj}(A)p_{obj}(B)$, then if the conditional subjective degrees of belief are stable in the sense of (18), then (assuming the Abstract Principal Principle) one has

$$
\begin{aligned}
& p_{subj}(h(A) \cap h(B) | \ulcorner p_{obj}(A) = r \urcorner \cap \ulcorner p_{obj}(B) = s \urcorner \cap \ulcorner p_{obj}(A \cap B) = t \urcorner) && (19) \\
&= p_{subj}(h(A \cap B) | \ulcorner p_{obj}(A) = r \urcorner \cap \ulcorner p_{obj}(B) = s \urcorner \cap \ulcorner p_{obj}(A \cap B) = t \urcorner) \\
&= p_{subj}(h(A \cap B) | \ulcorner p_{obj}(A \cap B) = t \urcorner) \\
&= p_{obj}(A \cap B) \\
&= p_{obj}(A)p_{obj}(B) \\
&= p_{subj}(h(A) | \ulcorner p_{obj}(A) = r \urcorner) p_{subj}(h(B) | \ulcorner p_{obj}(B) = s \urcorner) \\
&= p_{subj}(h(A) | \ulcorner p_{obj}(A) = r \urcorner \cap \ulcorner p_{obj}(B) = s \urcorner \cap \ulcorner p_{obj}(A \cap B) = t \urcorner) && (20) \\
&\cdot p_{subj}(h(B) | \ulcorner p_{obj}(A) = r \urcorner \cap \ulcorner p_{obj}(B) = s \urcorner \cap \ulcorner p_{obj}(A \cap B) = t \urcorner) && (21)
\end{aligned}
$$

Equations (19) and (20)-(21) mean that if the conditional subjective degrees of belief are stable, then, if A and B are objectively independent, then they (their isomorphic images $h(A), h(B)$) are also subjectively independent: they are independent also with respect to the probability measure that represents *conditional* subjective degrees of belief, where the condition is that the agent knows the objective probabilities of *all* of A, B and $(A \cap B)$. In other words, in this case the conditional subjective degrees of beliefs properly reflect the objective independence relations of random events – they are *independence-faithful*. To put this negatively: if a subjective probability measure satisfying the Abstract Principal Principle is *not* stable, then, although the agent's degrees of belief are equal to the objective probabilities of individual random events after a single conditionalization on the values of the objective probabilities of these individual events, these (unstable) individually conditionalized subjective probability measures do *not* necessarily reflect the objective independence relations between the random events – stable conditionalized subjective probabilities reflect the objective probabilities more faithfully than unstable ones. Note that for the subjective degrees of belief to satisfy the independence-faithfulness condition expressed by eqs. (19) and (20)-(21), it is sufficient that stability (18) only holds for the restricted set of elements B in the Boolean subalgebra $\mathcal{S}_{obj}^{A,ind}$ of $\mathcal{S}_{obj}$ generated by the elements in $\mathcal{S}_{obj}$ that are independent of A with respect to p_{obj}.

All this motivates to amend the Abstract Principal Principle by requiring stability of the subjective probabilities and define a "Stable Abstract Principal Principle":

Stable Abstract Principal Principle The subjective probabilities $p_{subj}(A)$ are related to the objective probabilities $p_{obj}(A)$ as

$$p_{subj}(A | \ulcorner p_{obj}(A) = r \urcorner) = p_{obj}(A) \tag{22}$$

Furthermore, the subjective probability function is *stable* in the sense that the following holds:

$$p_{subj}\big(h(A) | \ulcorner p_{obj}(A) = r \urcorner\big) = p_{subj}\big(h(A) | \ulcorner p_{obj}(A) = r \urcorner \cap \ulcorner p_{obj}(B) = s \urcorner\big) \qquad (\forall B \in \mathcal{S}) \tag{23}$$

If the subjective probability function is only *independence-stable* in the sense that (23) above holds for all $B \in \mathcal{S}_{obj}^{A,ind}$, then the corresponding Stable Abstract Principal Principle is called the *Independence-Stable* Abstract Principal Principle.

The next section raises the problem of the consistency of the Stable Abstract Principal Principle.

6 Is the Stable Abstract Principal Principle strongly consistent?

Definition 4. The Stable Abstract Principal Principle is defined to be *weakly consistent* if it is weakly consistent in the sense of Definition 1 and the subjective probability function p_{subj} is stable: it satisfies condition (23). The *Independence-Stable* Abstract Principal Principle is defined to be weakly consistent if it is weakly consistent in the sense of Definition 1 and the subjective probability function p_{subj} is *independence-stable*: it satisfies (23) for all $B \in \mathcal{S}_{obj}^{A,ind}$.

Thus the problem of weak consistency of the Stable Abstract Principal Principle emerges, and we have

Proposition 3. *The Stable Abstract Principal Principle is weakly consistent.*

The above proposition entails in particular

Proposition 4. *The Independence-Stable Abstract Principal Principle is weakly consistent.*

We prove Proposition 3 in the Appendix. The proof is based on the product extension technique that was used to show the weak consistency of the Abstract Principal Principle in the category of probability spaces with a finite Boolean algebra, and so the proof reveals the same weakness of the notion of weak consistency of the Stable Abstract Principal Principle that we pointed out earlier in connection with the Abstract Principal Principle: In the situation represented by the product extension case, the agent's degrees of beliefs $p_{subj}(A)$ are equal to the objective probabilities $p_{obj}(A)$ *without* any conditionalizing, and conditionalizing them on the independent propositions stating the values of the objective probabilities does not change the already correct degrees of belief. Clearly, this is a very exceptional situation however, and it is more realistic to assume that the agent's degrees of belief are only equal to the objective probabilities *after* conditionalizing them on knowing the values of objective probabilities but *not* before. One would like to know if the (Independence-)Stable Abstract Principal Principle is consistent under these more realistic circumstances. The definition of weak consistency of the (Independence-)Stable Abstract Principal Principle does not say anything about this more stringent consistency however; thus it is desirable to strengthen it in the manner specified in the next definition.

Definition 5. The Stable Abstract Principal Principle is defined to be *strongly consistent* if it is strongly consistent in the sense of Definition 2 and the subjective probability function p_{subj} is stable. Explicitly:
The Stable Abstract Principal Principle is strongly consistent if the following hold: Given any probability space $(X_{obj}, \mathcal{S}_{obj}, p_{obj})$ and another probability measure p^0_{subj} on $\mathcal{S}_{obj}$, there exists a probability space $(X_{subj}, \mathcal{S}_{subj}, p_{subj})$ and a Boolean algebra embedding h of $\mathcal{S}_{obj}$ into $\mathcal{S}_{subj}$ such that

(i) For every $A \in \mathcal{S}_{obj}$ there exists an $A' \in \mathcal{S}_{subj}$ with the property

$$p_{subj}(h(A)|A') = p_{obj}(A) \tag{24}$$

(ii) If $A, B \in \mathcal{S}_{obj}$ and $A \neq B$ then $A' \neq B'$.

(iii) The probability space $(X_{subj}, \mathcal{S}_{subj}, p_{subj})$ is an extension of the probability space $(X_{obj}, \mathcal{S}_{obj}, p^0_{subj})$ with respect to h; i.e. we have

$$p_{subj}(h(A)) = p^0_{subj}(A) \qquad A \in \mathcal{S}_{obj} \tag{25}$$

(iv) For all $B \in \mathcal{S}$ we have

$$p_{subj}(h(A)|A') = p_{subj}\big(h(A)|A' \cap B'\big) \tag{26}$$

The *Independence-Stable* Abstract Principal principle is strongly consistent if (i)-(iii) above holds, and instead of (iv), we have

(iv')

$$p_{subj}(h(A)|A') = p_{subj}\big(h(A)|A' \cap B'\big) \qquad \forall B \in \mathcal{S}^{A,ind}_{obj} \tag{27}$$

Problem 1. Is the (Independence-)Stable Abstract Principal Principle strongly consistent?

The problem of strong consistency of both the Stable and of the Independence-Stable Abstract Principal Principle remain open, even in the category of probability spaces with a Boolean algebra having a finite number of elements: An explicit counterexample detailed in the Appendix (Proposition 7) shows that the method used in proving the strong consistency of the Abstract Principal Principle does *not* prove the strong consistency of the *Stable* Abstract Principal Principle. Specifically, one can show by an explicit calculation that if a Bayesian agent is in the very simple epistemic situation of having to form degrees of beliefs about the two events occurring in coin flipping, which is described by the simplest non-trivial Boolean algebra $\mathcal{S}^4_{obj}$ formed by the four events $\{\emptyset, A(Head), A^{\perp}(Tail), I\}$ with objective probabilities $p_{obj}(A) = \frac{1}{5}$ and $p_{obj}(A^{\perp}) = \frac{4}{5}$, then, if a Bayesian agent's prior subjective degrees of beliefs are $p^0_{subj}(A) = \frac{1}{3}$ and $p^0_{subj}(A^{\perp}) = \frac{2}{3}$, then the subjective degree of belief p_{subj} on the Boolean algebra $\mathcal{S}_{subj}$ constructed using the method in the proof of Proposition 2 will *not* be stable. Note that we do *not* claim, that in such an epistemic situation it is not possible for the Bayesian agent to form a $\mathcal{S}_{subj}$ and extend p^0_{subj} to $\mathcal{S}_{subj}$ in such a way that the Abstract Principal Principle holds for p_{subj} possibly together with stability. We only claim that it is not possible to do this in the particular way that proves in general the strong consistency of the Abstract Principal Principle; hence the strong consistency of the Stable Abstract Principal Principle remains an open problem.

In the next two sections we argue that proving either that the (Independence-)Stable Principal Principle is strongly consistent or that the strong consistency does not hold has ramifications both for the original Principal Principle and for Bayesianism in general.

7 Can Bayesian agents always be rational?

Call a Bayesian agent *omniscient* if he never has to adjust his subjective degrees of belief because his prior degrees of belief about random events in $\mathcal{S}_{obj}$ coincide with the objective probabilities p_{obj} on $\mathcal{S}_{obj}$ *in all epistemic situations* without any conditionalization: no matter what the set of random events $\mathcal{S}_{obj}$ and their objective probabilities p_{obj} are, facing $(\mathcal{S}_{obj}, p_{obj})$, the agent has prior degrees of belief p^0_{subj} on $\mathcal{S}_{obj}$ such that

$$p^0_{subj}(A) = p_{obj}(A) \qquad A \in \mathcal{S}_{obj} \tag{28}$$

The weak consistency of the Stable Abstract Principal Principle shows that such an omniscient agent can always be a *rational* Bayesian agent in the sense that (i) he can always form a logically well-behaving set of propositions which is closed with respect to Boolean operations and which contains *both* the propositions stating that events in $\mathcal{S}$ happen *and* the propositions about the values of their objective probability; furthermore (ii) he can always have consistent degrees of belief as probabilities about *all* these propositions in such a way that (iii) if he conditionalizes his already correct prior degrees of belief on the values of any of the objective probabilities, his correct degrees of belief do not change. In short: weak consistency of the Stable Abstract Principal Principle ensures that an *omniscient* Bayesian agent can in principle be *always* rational.

We regard the strong consistency of the Stable Abstract Principal Principle necessary for a *non-omniscient* (hence more realistic) Bayesian agent to be able to have *rational* degrees of belief *in all epistemic situations*. For if the strong consistency of the Stable Abstract Principal Principle does *not* hold, then there exist in principle epistemic situations the Bayesian agent can find himself in, in which at least one of the following (i)-(iii) *cannot* be maintained:

(i) Degrees of beliefs of the agent are represented by probability measures satisfying the usual axioms of measure theoretic probability.

(ii) The agent's prior degrees of beliefs differ from the objective probabilities but by learning the correct objective probabilities the agent can adjust his degrees of beliefs using Bayesian conditionalization so that they become equal to the objective probabilities.

(iii) The adjusted degrees of beliefs are stable: by learning additional truths about objective probabilities and re-conditionalizing his correct degrees of belief on them, the agent is not loosing his already correct degrees of belief.

If not even the *Independence-Stable* Abstract Principal Principle can be proved to be strongly consistent, then (iii) above can be replaced with

(iii') The adjusted degrees of beliefs are stable: by learning additional truths about objective probabilities of *objectively independent events* and re-conditionalizing his correct degrees of belief on them, the agent is not loosing his already correct degrees of belief (which also entails that the agent's conditioned degrees of belief reflect the objective independence properties of random events).

Since strong consistency of both the Stable and of the Independence-Stable Abstract Principal Principle remain open problems, we have to conclude that at this point we do not know for certain whether non-omniscient Bayesian agents can *always* be rational in principle.

8 Comments on the Principal Principle involving chances

The idea of the Abstract Principal Principle can be traced back to Reichenbach's "Straight Rule" of induction connecting subjective and objective probabilities (the paper [3] gives a comprehensive review of Reichenbach's inductive logic and the role of the straight rule in it). The first claim about a possible inconsistency of the Straight Rule seems to be Miller's Paradox [13]; since Miller's work, the Straight Rule in the form of equation (1) is also called "Miller's Principle" and "Minimal Principle" [17], [14]. The inconsistency claim by Miller did not have anything to do with the type of consistency investigated in the present paper and Miller's Paradox was shown to be a pseudo-paradox resulting from the ambiguity of the formulation of the Straight Rule [12], [2], [7].

Lewis [9] introduced the term "Principal Principle" to refer to the specific principle that links subjective beliefs to chances in the manner expressed by (1): In the context of the Principal Principle $p_{subj}(A)$ is called the "credence", $Cr_t(A)$, of the agent in event A at time t, $p_{obj}(A)$ is the chance $Ch_t(A)$ of the event A at time t, and the Principal Principle is the stipulation that credences and chances are related as

$$Cr_t(A|\ulcorner Ch_t(A) = r\urcorner \;\&\; E) = Ch_t(A) = r \tag{29}$$

where E is any *admissible* evidence the agent has at time t in addition to knowing the value of the chance of A.

Lewis himself saw a consistency problem in connection with the Principal Principle (he called it the "Big Bad Bug"): If A is an event in the future of t that has a non-zero chance $r > 0$ of happening at that later time but we have knowledge E about the future that entails that A will in fact not happen, $E \subset A^{\perp}$, then substituting this E into (29) leads to contradiction if $r > 0$. Such an A is called an "unactualized future that undermines present chances" – hence the phrase "undermining" to refer to this situation. Since certain metaphysically motivated arguments based on a Humean understanding of chance led Lewis to think that one is forced to admit such an evidence E, he tried to "debug" the Principal Principle [10]; the same sort of debugging was proposed simultaneously by Hall [5] and Thau [16]. A number of other debugging attempts and modifications have followed [11], [6], [8], [15], and to date no consensus has emerged as to which of the debugged versions of Lewis original Principal Principle is tenable: Vranas claims [17] that there was no need for a debugging in the first place; Briggs [1] argues that none of the modified principles work; Pettigrew [14] provides a systematic framework that allows in principle to choose the correct Principal Principle depending on one's metaphysical concept of chance.

The relevance of the strong consistency of the (Independence-)Stable Abstract Principal Principle for the original Principal Principle and its debugged versions should now be clear: If admissibility of evidence E is defined in such a way that E propositions about the values of chances of events are admitted, then the consistency of the corresponding Principal Principle becomes an open question because the problem of strong consistency of the (Independence-)Stable Abstract Principal Principle is open. Allowing this kind of evidence seems common: Lewis himself regarded admissible all propositions containing information that is "irrelevant" for the chance of A [9][p. 91], these should include propositions about values of chances of events that are independent of A with respect to the probability measure describing their chances. Ismael's New Proposal admits any proposition about events in the backward light cone (causal past) of event A as admissible evidence [8][p. 296]; presumably propositions about chances of these events belong to the admissible class. Hoefer's informal specification of admissibility [6][p. 553] also seems to admit propositions stating values of chances of events as admissible evidence. Consequently, the consistency of *all* these Principal Principles is an open question. This kind of consistency has nothing to do with any metaphysics about chances or with the concept of natural laws that one may have in the background of the Principal Principle: This consistency expresses a simple but fundamental compatibility of the Principal Principle with the basic structure of probability theory. Without proving this consistency it is not clear whether the Principal Principle can be formulated at all meaningfully in terms of probability theory involving the Bayes rule.

9 Appendix

9.1 Proof of weak consistency of the Stable Abstract Principal Principle (Proposition 3)

The statement of weak consistency of the Stable Principal Principle follows from Proposition 5 below if we make the following identifications:

- $(X_{obj}, \mathcal{S}_{obj}, p_{obj}) \leftrightarrow (X, \mathcal{S}, p)$
- $(X_{subj}, \mathcal{S}_{subj}, p_{subj}) \leftrightarrow (X', \mathcal{S}', p')$

Proposition 5. *Let $(X, \mathcal{S}, p)$ be a probability space. Then there exists an extension $(X', \mathcal{S}', p')$ of $(X, \mathcal{S}, p)$ with respect to a Boolean algebra homomorphism $h\colon \mathcal{S} \to \mathcal{S}'$ such that*

(i) For all $A \in \mathcal{S}$ there is $A' \in \mathcal{S}'$ such that

$$p'\big(h(A)|A'\big) = p(A)$$

(ii) $A \neq B$ implies $A' \neq B'$

(iii)

$$p'(h(A)|A') = p'\big(h(A)|A' \cap B'\big) \qquad (\forall B' \in \mathcal{S}) \tag{30}$$

Proof. Let $(X, \mathcal{S}, p)$ be a probability space and Y_0 be a set disjoint from $\mathcal{S}$ and having the same cardinality as the cardinality of $\mathcal{S}$. We can think of Y_0 as having elements y_A labeled by elements $A \in \mathcal{S}$. Consider the set

$$Y \doteq Y_0 \cup \{y\} = \{y_A : A \in \mathcal{S}\} \cup \{y\}$$

where y is an auxiliary element different from every y_A. Take the powerset $\mathcal{P}(Y)$ and let q be any probability measure on $\mathcal{P}(Y)$ such that $q(\{y\}) \neq 0$. Then $(Y, \mathcal{P}(Y), q)$ is a probability space and we can form the product space

$$(X', \mathcal{S}', p') = (X \times Y, \mathcal{S} \otimes \mathcal{P}(Y), p \times q)$$

with $p' = (p \times q)$ being the product measure on $\mathcal{S} \otimes \mathcal{P}(Y)$. Recall that by definition of the product measure, for sets of the form $A \times B$ ($A \in \mathcal{S}$, $B \in \mathcal{P}(Y)$) we have

$$p'(A \times B) = p(A)q(B)$$

The map $h : \mathcal{S} \to \mathcal{S}'$ defined by $h(A) \doteq A \times Y$ is an injective, measure preserving Boolean algebra embedding. For each $A \in \mathcal{S}$ put

$$A' \doteq X \times \{y_A, y\}$$

It is clear that (ii) in the proposition holds for A', B' so defined.
To see (i) one can compute:

$$\begin{aligned} p'(h(A)|A') &= \frac{p'(h(A) \cap A')}{p'(A')} = \frac{p'\big((A \times Y) \cap (X \times \{y_A, y\})\big)}{p'(X \times \{y_A, y\})} \\ &= \frac{p'\big(A \times \{y_A, y\}\big)}{p'\big(X \times \{y_A, y\}\big)} = \frac{p(A) \cdot q(\{y_A, y\})}{p(X) \cdot q(\{y_A, y\})} = p(A) \end{aligned}$$

Condition (iii) also holds because

$$\begin{aligned} p'(h(A)|A' \cap B') &= \frac{p'(h(A) \cap A' \cap B')}{p'(A' \cap B')} = \frac{p'\big((A \times Y) \cap (X \times \{y_A, y\}) \cap (X \times \{y_B, y\})\big)}{p'((X \times \{y_A, y\}) \cap (X \times \{y_B, y\}))} \\ &= \frac{p'\big(A \times \{y\}\big)}{p'\big(X \times \{y\}\big)} = \frac{p(A) \cdot q(\{y\})}{p(X) \cdot q(\{y\})} = p(A) \end{aligned}$$

□

9.2 Proof of strong consistency of the Abstract Principal Principle in the category of probability spaces with a finite Boolean algebra (Proposition 2)

The statement of strong consistency of the Stable Principal Principle in the category of probability spaces follows from Proposition 6 below if we make the following identifications:

- $(X_{obj}, \mathcal{S}_{obj}, p_{obj}) \leftrightarrow (X, \mathcal{S}, \hat{p})$
- $(X_{obj}, \mathcal{S}_{obj}, p^0_{subj}) \leftrightarrow (X, \mathcal{S}, p)$
- $(X_{subj}, \mathcal{S}_{subj}, p_{subj}) \leftrightarrow (X', \mathcal{S}', p')$

Proposition 6. *Let $(X, \mathcal{S}, p)$ be a probability space with $\mathcal{S}$ having $n < \infty$ elements and let $\hat{p}$ be another probability measure on $\mathcal{S}$ such that p is non-extreme with respect to $\hat{p}$ (Definition 3): $p(A) = 1$ implies $\hat{p}(A) = 1$ and $p(A) = 0$ implies $\hat{p}(A) = 0$. Then there exists an extension $(X', \mathcal{S}', p')$ of $(X, \mathcal{S}, p)$ with respect to the embedding $h\colon \mathcal{S} \to \mathcal{S}'$ having the following properties:*

(i) For all $A \in \mathcal{S}$ there is $A' \in \mathcal{S}'$ such that

$$p'\big(h(A)|A'\big) = \hat{p}(A)$$

(ii) $A \neq B$ implies $A' \neq B'$

Proof. The proof consist of two steps. In the first step we choose an arbitrary element $A \in \mathcal{S}$ and construct an extension of $(X^*, \mathcal{S}^*, p^*)$ with respect to an embedding h^* in such a manner that in this extension this particular event A has a pair $A' = A^*$ with the required properties. In step 2 we then repeat this step $n-1$ times, choosing each time another element from $\mathcal{S}$ until we exhaust $\mathcal{S}$ and obtain the extension $(X', \mathcal{S}', p')$ of $(X, \mathcal{S}, p)$.

STEP 1. Take any $A \in \mathcal{S}$. We wish to construct a space $(X^*, \mathcal{S}^*, p^*)$ and a function $h^* : \mathcal{S} \to \mathcal{S}^*$ such that

- $h^* : (\mathcal{S}, p) \to (\mathcal{S}^*, p^*)$ is a measure preserving, injective Boolean algebra homomorphism.
- There is $A^* \in \mathcal{S}^*$ such that $p^*\big(h^*(A)|A^*\big) = \hat{p}(A)$.

Take two disjoint copies of $(X, \mathcal{S})$, let these copies be $(X_1, \mathcal{S}_1)$ and $(X_2, \mathcal{S}_2)$ and fix the algebra isomorphisms $h_1 : (X, \mathcal{S}) \to (X_1, \mathcal{S}_1)$ and $h_2 : (X, \mathcal{S}) \to (X, \mathcal{S}_2)$. Put $X^* = X_1 \cup X_2$ and define

$$\mathcal{S}^* = \big\{h_1(A) \cup h_2(B) : A, B \in \mathcal{S}\big\} \tag{31}$$

It is a routine task to verify that $(X^*, \mathcal{S}^*)$ is a measurable space, i.e. $\mathcal{S}^*$ is a Boolean algebra of subsets of X^* with respect to the usual set theoretical operations $\cup$, $\cap$, $\perp$. Define the map $h^* : \mathcal{S} \to \mathcal{S}^*$ as follows

$$h^*(A) = h_1(A) \cup h_2(A) \qquad\qquad A \in \mathcal{S} \tag{32}$$

We claim that h^* is a homomorphism between $\mathcal{S}$ and $\mathcal{S}^*$. Indeed: take any $A, B \in \mathcal{S}$ and observe that

$$\begin{aligned}
&h^*(A \cup B) = h_1(A \cup B) \cup h_2(A \cup B) = h_1(A) \cup h_1(B) \cup h_2(A) \cup h_2(B) = h^*(A) \cup h^*(B) \\
&h^*(X \setminus A) = h_1(X \setminus A) \cup h_2(X \setminus A) = \big(X_1 \setminus h_1(A)\big) \cup \big(X_2 \setminus h_2(A)\big) = \\
&= (X_1 \cup X_2) \setminus (h_1(A) \cup h_2(A)) = X^* \setminus h^*(A)
\end{aligned}$$

Let $0 \leq \alpha \leq 1$ be any real number and define p^* by

$$p^*\big(h_1(A) \cup h_2(B)\big) \doteq \alpha \cdot p(A) + (1 - \alpha) \cdot p(B) \tag{33}$$

Note that for each $A \in \mathcal{S}$ we have, by definition

$$p^*\big(h^*(A)\big) = \alpha \cdot p(A) + (1 - \alpha) \cdot p(A) = p(A) \tag{34}$$

Consequently, $h^* : (\mathcal{S}, p) \to (\mathcal{S}^*, p^*)$ is a measure preserving, injective Boolean algebra homomorphism.

For any fixed $A \in \mathcal{S}$ define A^* by

$$A^* \doteq h_1(A) \cup h_2(A^\perp) \tag{35}$$

Our aim now is to choose α in such a way that the following is true:

$$p^*(h^*(A)|A^*) = \hat{p}(A) \tag{36}$$

Some basic algebra shows that

$$\begin{aligned}
p^*(h^*(A)|A^*) &= \frac{p^*(h^*(A) \cap A^*)}{p^*(A^*)} \\
&= \frac{p^*\Big(\big(h_1(A) \cup h_2(A)\big) \cap \big(h_1(A) \cup h_2(A^\perp)\big)\Big)}{p^*(h_1(A) \cup h_2(A^\perp))} \\
&= \frac{p^*(h_1(A) \cup \emptyset)}{p^*(h_1(A) \cup h_2(A^\perp))} \\
&= \frac{\alpha \cdot p(A) + (1 - \alpha) \cdot p(\emptyset)}{\alpha \cdot p(A) + (1 - \alpha) \cdot p(A^\perp)} \\
&= \frac{\alpha \cdot p(A)}{\alpha \cdot p(A) + (1 - \alpha) \cdot (1 - p(A))}
\end{aligned}$$

This means that we have to carefully choose α to guarantee

$$\frac{\alpha \cdot p(A)}{\alpha \cdot p(A) + (1 - \alpha) \cdot (1 - p(A))} = \hat{p}(A) \tag{37}$$

By our assumptions, if $p(A) = 1$ then $\hat{p}(A) = 1$ and thus any $\alpha \neq 0$ makes (37) true. Similarly, if $p(A) = 0$, then $\hat{p}(A) = 0$, which means that any $\alpha \neq 1$ will do. Also, if $\hat{p}(A) = 0$, then $\alpha = 0$ will do. Therefore we may assume $0 < p(A) < 1$ and $0 < \hat{p}(A) \leq 1$. By re-ordering equation (37) and using the notation $p = p(A)$, $r = \hat{p}(A)$ we obtain the equation

$$\alpha = \frac{rp - r}{rp - r + pr - p} \tag{38}$$

In order to guarantee (37) we only have to show that α in equation (38) is between 0 and 1. To do so, observe that since $0 < p < 1$ and $0 < r \leq 1$ we have $rp < r$ and $pr \leq p$. This means that both the numerator and the denominator of the fraction in (38) is negative, whence α is positive. On the other hand, we have

$$\begin{aligned} 0 &\geq pr - p \\ rp - r &\geq rp - r + pr - p \\ \frac{rp - r}{rp - r + pr - p} &\leq 1 \end{aligned}$$

Thus we proved that $0 \leq \alpha \leq 1$ can always be chosen so that equation (36) holds.
STEP 2. We obtain $(X', \mathcal{S}', p')$ by iterating Step 1. above. Let $A_1, \ldots, A_n$ be a one-to-one enumeration of $\mathcal{S}$. Applying step one to A_1 one finds a space $(X_1, \mathcal{S}_1, p_1)$, an event $A_1^* \in \mathcal{S}_1$ and an embedding h_1

$$(X, \mathcal{S}, p) \xrightarrow{h_1} (X_1, \mathcal{S}_1, p_1),$$

such that $p_1\big(h_1(A_1)|A_1^*\big) = \hat{p}(A_1)$. In a similar manner, step one applied to $h_1(A_2)$ gives a space $(X_2, \mathcal{S}_2, p_2)$, an element $A_2^* \in \mathcal{S}_2$ and an embedding h_2

$$(X, \mathcal{S}, p) \xrightarrow{h_1} (X_1, \mathcal{S}_1, p_1) \xrightarrow{h_2} (X_2, \mathcal{S}_2, p_2)$$

so that $p_2\big(h_2 h_1(A_2)|A_2^*\big) = \hat{p}(A_2)$. Now, one has to verify that the second extension does not destroy the result of the first one, that is to say, $p_2\big(h_2 h_1(A_1)|h_2(A_1^*)\big)$ should remain equal to $\hat{p}(A_1)$. But as h_2 is an embedding, it preserves measure, thus it preserves conditional measure as well, meaning that

$$p_2\big(h_2 h_1(A_1)|h_2(A_1^*)\big) = p_1\big(h_1(A_1)|A_1^*\big) = \hat{p}(A_1).$$

Continuing in this line, after n steps we get elements $A_i^* \in \mathcal{S}_i$ and a chain of extensions

$$(X, \mathcal{S}, p) \xrightarrow{h_1} (X_1, \mathcal{S}_1, p_1) \xrightarrow{h_2} (X_2, \mathcal{S}_2, p_2) \rightarrow \cdots \xrightarrow{h_n} (X_n, \mathcal{S}_n, p_n)$$

such that

$$p_n\big(h_n \cdots h_2 h_1(A_i)|h_n \cdots h_{i+1}(A_i^*)\big) = \hat{p}(A_i)$$

holds for all A_i. Therefore we can complete the proof by letting

$$\begin{aligned} (X', \mathcal{S}', p') &= (X_n, \mathcal{S}_n, p_n) \\ h &= h_n h_{n-1} \cdots h_1 \\ A_i' &= h_n \cdots h_{i+1}(A_i^*) \end{aligned}$$

(where h_{n+1} is the identity, of course). □

Proposition 7. *There exists a probability space $(X, \mathcal{S}, \hat{p})$ and a probability measure p on $\mathcal{S}$ such that the probability measure p' in the probability space $(X', \mathcal{S}', p')$ constructed in the proof of Proposition 2 will not be stable, i.e. for some $A, B \in \mathcal{S}$ we have:*

$$p'(h(A)|A') \neq p'(h(A)|A' \cap B') \tag{39}$$

Recall (see the statement of Proposition 2) that $(X, \mathcal{S}, \hat{p})$ is the probability space describing the objective random events and their probabilities, $(X', \mathcal{S}', p')$ stand for $(X_{subj}, \mathcal{S}_{subj}, p_{subj})$ with p' being the extension of the agent's prior probability p on $\mathcal{S}$.

Proof. Let $\mathcal{S} = \mathcal{S}^4$ be the Boolean algebra having four elements $\{\emptyset, A, B, I\}$ (clearly: $B = A^\perp$) and let $\hat{p}$ be a probability measure on $\mathcal{S}^4$. Assume that the probability p is $p(A) = x$ and $p(B) = p(A^\perp) = 1 - x$ for some real number $x \in [0, 1]$. Since $\mathcal{S}^4$ has only two non-trivial elements, constructing the space $(X', \mathcal{S}', p')$ that extends $(X, \mathcal{S}, p)$ in the way detailed in the proof of Proposition 6 is carried out in two steps: First one constructs $(\mathcal{S}^*, p^*)$ with the Boolean algebra homomorphism $h^* \colon \mathcal{S} \to \mathcal{S}^*$ in such a way that equations (33), (34), (36) and (37) hold with a suitable α. In the second step one constructs the extension $(\mathcal{S}^{**}, p^{**})$ of $(\mathcal{S}^*, p^*)$ with the Boolean algebra homomorphism $h^{**} \colon \mathcal{S}^* \to \mathcal{S}^{**}$ in such a way that the analogues of equations (33), (34), (36) and (37) hold, now replacing A with $h^*(B)$ and α with a suitable β. Following the notation in the proof of Proposition 6, and in particular eq. (35), we can write:

$$h = h^{**} \circ h^* \tag{40}$$
$$A' = h^{**}(A^*) \tag{41}$$
$$B' = \big(h(B)\big)^* \tag{42}$$

the equations expressing the Abstract Principal Principle in connection with events A and B are

$$p'\big(h(A)|A'\big) = \frac{\alpha x}{\alpha x + (1-\alpha)(1-x)} \tag{43}$$

$$p'\big(h(B)|B'\big) = \frac{\beta(1-x)}{\beta(1-x) + (1-\beta)x} \tag{44}$$

One can now compute explicitly the subjective probability $p'\big(h(A)|A'\cap B'\big)$ of $h(A)$ in the probability space $(X', \mathcal{S}', p')$ after a *second* conditionalization on the value of the objective probability probability of B. The computation yields

$$p'\big(h(A)|A'\cap B'\big) = \frac{(1-\beta)\alpha x}{\beta(1-\alpha)(1-x) + (1-\beta)\alpha x} \tag{45}$$

Note that since $p(h(A)|A') = \hat{p}(A)$ and $p'(h(B)|B') = \hat{p}(B) = \hat{p}(A^\perp)$, we have

$$p(h(A)|A') + p(h(B)|B') = 1 \tag{46}$$

hence if we take $x = \alpha = \frac{1}{3}$, then we get

$$p(h(A)|A') = \frac{1}{5}$$

$$p(h(B)|B') = \frac{4}{5} \Longrightarrow \beta = \frac{2}{3}$$

$$p(h(A)|A'\cap B') = \frac{1}{9}$$

Therefore $p(h(A)|A') \neq p(h(A)|A'\cap B')$.

□

References

[1] R. Briggs. The anatomy of the Big, Bad Bug. *Noûs*, 43:428–449, 2009.

[2] J. Bub and M. Radner. Miller's paradox of information. *The British Journal for the Philosophy of Science*, 19:63–67, 1968.

[3] F. Eberhardt and C. Glymour. Hans Reichenbach's probability logic. In D.M. Gabbay, S. Hartmann, and J. Woods, editors, *Handbook of the History of Logic*, volume 10. Inductive Logic. Elsevier, 2009.

[4] Z. Gyenis and M. Rédei. How much can a Bayesian agent learn in principle? *Manuscript*, 2013. In preparation.

[5] N. Hall. Correcting the guide to objective chance. *Mind*, 103:505–518, 1994.

[6] C. Hoefer. The third way on objective probability: A sceptic's guide to objective chance. *Mind*, 116:449–596, 2007.

[7] C. Howson and G. Oddie. Miller's so-called paradox of information. *The British Journal for the Philosophy of Science*, 30:253–278, 1979.

[8] J. Ismael. Raid! Correcting the Big Bad Bug. *Noûs*, 42:292–307, 2008.

[9] D. Lewis. A subjectivist's guide to objective chance. In *Philosophical Papers, vol. II*, pages 83–132. Oxford University Press, Oxford, 1986.

[10] D. Lewis. Humean supervenience debugged. *Mind*, 103:473–490, 1994.

[11] B. Loewer. David Lewis' Humean theory of objective chance. *Philosophy of Science*, 71:115–1125, 2004.

[12] J. Mackie. Miller's so-called paradox of information. *The British Journal for the Philosophy of Science*, 14:144–147, 1966.

[13] D. Miller. A paradox of information. *The British Journal for the Philosophy of Science*, 17:59–61, 1966.

[14] R. Pettigrew. Accuracy, chance and the Principal Principle. *Philosophical Review*, 121:241–275, 2012.

[15] J.T. Roberts. Undermining undermined: Why Humean supervenience never needed to be debugged. *Philosophy of Science*, 68:S98–S108, 2001. Proceedings of the 2000 Biennial Meeting of the Philosophy of Science Association. Part I Contributed Papers.

[16] M. Thau. Undermining and admissibility. *Mind*, 103:491–504, 1994.

[17] P.B.M. Vranas. Have your cake and eat it too: The Old Principal Principle reconciled with the New. *Philosophy and Phenomenological Research*, LXIX:368–382, 2004.