Economics & Philosophy (2020), **36**, 127–147
doi:10.1017/S0266267118000536

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

Countable additivity, idealization, and conceptual realism

Yang Liu

Faculty of Philosophy, University of Cambridge, Cambridge CB3 9DA, UK
Email: liu@yliu.net

(Received 26 April 2018; revised 20 August 2018; accepted 23 August 2018; first published online 28 February 2019)

Abstract

This paper addresses the issue of finite versus countable additivity in Bayesian probability and decision theory – in particular, Savage's theory of subjective expected utility and personal probability. I show that Savage's reason for not requiring countable additivity in his theory is inconclusive. The assessment leads to an analysis of various highly idealized assumptions commonly adopted in Bayesian theory, where I argue that a healthy dose of, what I call, conceptual realism is often helpful in understanding the interpretational value of sophisticated mathematical structures employed in applied sciences like decision theory. In the last part, I introduce countable additivity into Savage's theory and explore some technical properties in relation to other axioms of the system.

Keywords: countable additivity; foundations of probability; Bayesian decision theory; idealization; conceptual realism

1. Introduction

One recurring topic in philosophy of probability concerns the additivity condition of probability measures as to whether a numerical probability function should be *finitely* or *countably additive* (henceforth FA and CA respectively for short[1]). The issue is particularly controversial within Bayesian subjective decision and probability theory, where probabilities are seen as measures of agents' degrees of beliefs, which are interpreted within a general framework of rational decision making. As one of the founders of Bayesian subjectivist theory, de Finetti famously rejected CA because (as I will explain below) he takes that CA is in tension with the subjectivist interpretation of probability. De Finetti's view, however, was contested among philosophers.

The debate on FA versus CA in Bayesian models often operates on two fronts. First, there is the concern of mathematical consequences as a result of assuming either FA or CA. Proponents of CA often refer to a pragmatic (or even a sociological)

[1]In what follows 'FA' will be used ambiguously for both 'finitely additive', and 'finite additivity', similarly for 'CA'. The use of singular 'they' will also be adopted where appropriate.

© Cambridge University Press 2019. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

https://doi.org/10.1017/S0266267118000536 Published online by Cambridge University Press

point that it is a common practice in mathematics that probability measures are taken to be CA ever since the first axiomatization of probability theory by Kolmogorov. Among the advantages of assuming CA is that it allows the subjective interpretation of probability to keep step with the standard and well-established mathematical theory of probability. However, alternative ways of establishing a rich theory of probability based on FA are also possible.[2] Hence different *mathematical reasons* are cited by both sides of the debate as bearing evidence either for or against the choice of FA or CA.

The second general concern is about the *conceptual* underpinning of different additivity conditions under the subjective interpretation of probability. As noted above, within the subjectivist framework, a theory of personal probability is embedded within a general theory of rational decision making, couched in terms of how agents' probabilistic beliefs coherently guide their actions. Such normative theories of actions are often based on a series of rationality postulates governing an agent's partial beliefs as well as their choice behaviours in decision situations. This then gives rise to the question as to how the choice of FA or CA – an otherwise unremarkable mathematical property of some additive function – is accounted for within the subjectivist theory. There is thus the problem of explaining why (and what) rules for rational beliefs and actions should always respect one additivity condition rather than the other.

Much has been written about de Finetti's reasons for rejecting CA on both fronts. I will refer to some of these arguments and related literature below. My focus here, however, is on the issue of additivity condition within Savage's theory of subjective expected utility (Savage 1954, 1972). The latter is widely seen as *the* paradigmatic system of subjective decision making, on which a classical theory of personal probability is based. Following de Finetti, Savage also cast out CA for probability measures derived in his system. One goal of this paper is to point out that the arguments enlisted by Savage were inconclusive. Accordingly, I want to explore ways of introducing CA into Savage's system, which I will pursue in the last part of this paper.

As we shall see, the discussion will touch upon various highly idealized assumptions on certain underlying logical and mathematical structures that are commonly adopted in Bayesian probability and decision theory. The broader philosophical aim of this paper is thus to provide an analysis of these assumptions, where I argue that a healthy dose of, what I call, *conceptual realism* is often helpful in understanding the interpretational values of certain sophisticated mathematics involved in applied sciences like Bayesian decision theory.

1.1. The measure problem

To begin, in a section titled 'Some mathematical details' Savage (1972) explained his take on additivity conditions in his subjectivist theory, he says,

> It is not usual to suppose, as has been done here, that *all* sets have a numerical probability, but rather a sufficiently rich class of sets do so, the remainder being considered unmeasurable . . . the theory being developed here does assume that probability is defined for *all events*, that is, for *all sets of states*, and it does not imply countable additivity, but only finite additivity . . . it is a little better

[2]Rao and Rao (1983), for instance, is a study of probability calculus based on FA.

https://doi.org/10.1017/S0266267118000536 Published online by Cambridge University Press

> not to assume countable additivity as a postulate, but rather as a special hypothesis in certain contexts. (Savage 1972: 40, emphasis added)

One main mathematical reason provided by Savage for not requiring CA is that there does not exist, it is said, a countably additive extension of the Lebesgue measure definable over *all* subsets of the unit interval (or the real line), whereas in the case of finitely additive measures, such an extension does exist. Since events are taken to be 'all sets of states' in his system (all subsets of the reals, $\mathcal{P}(\mathbb{R})$, in the case where the state space is the real line), CA is ruled out because of this claimed defect.

Savage's remarks refer to the basic problem of measure theory posed by Henri Lebesgue at the turn of the 20th century known as the *problem of measurability*.[3] Lebesgue himself developed a measure towards the solution to this problem. Unlike other attempts made around the same period (Bingham 2000), the measure developed by him, later known as the *Lebesgue measure*, was constructed in accordance with certain algebraic structure of sets of the reals. As seen, the measure problem would be solved if it could be shown that the Lebesgue measure satisfies all the measurability conditions (i.e. conditions (a)–(d) in fn 3).

The measure problem, however, was soon answered in the negative by Vitali (1905), who showed that, in the presence of the Axiom of Choice (AC), there exist sets of real numbers that are not (Lebesgue) measurable. This means that, with AC, Lebesgue's measure is definable only for a proper class of all subsets of the reals, the remainder being unmeasurable. Then a natural question to ask is whether there exists an 'extension' of the Lebesgue measure such that it not only agrees with the Lebesgue measure on all measurable sets, but is also definable for non-measurable ones. Call this the *revised measure problem*. This revised problem gives rise to a more general question as to whether or not there exists a real-valued measure on any infinite set.

To anticipate our discussion on subjective probabilities and Savage's reasons for relaxing the CA condition, let us reformulate the question in terms of probabilistic measures defined over some infinite set. Let S be a (countably or uncountably) infinite set, a *measure* on S is a non-negative real-valued function μ on $\mathcal{P}(S)$ such that

(i) μ is defined for all subsets of S;
(ii) $\mu(\varnothing) = 0$, $\mu(S) = 1$;
(iii) μ is countably additive (or σ-additive), that is, if $\{X_n\}_{n=1}^{\infty}$ is a collection of pairwisely disjoint bounded subsets of $\mathbb{R}$, then

$$\mu\left(\bigcup_{n=1}^{\infty} X_n\right) = \sum_{n=1}^{\infty} \mu(X_n). \tag{1}$$

[3]More precisely, in his 1902 thesis Lebesgue raised the following question about the real line: Does there exist a measure m such that (a) m associates with each bounded subset X of $\mathbb{R}$ a real number $m(X)$; (b) m is non-trivial, i.e. $m(X) \neq 0$ for all $X \neq \varnothing$; (c) m is translation-invariant, that is, for any $X \subseteq \mathbb{R}$ and any $r \in \mathbb{R}$, define $X + r := \{x + r \mid x \in X\}$, then $m(X) = m(X + r)$; and (d) m is σ-additive, that is, if $\{X_n\}_{n=1}^{\infty}$ is a collection of pairwise disjoint bounded subsets of $\mathbb{R}$, then $m(\bigcup_n X_n) = \sum_n m(X_n)$? See Hawkins (1975) for a detailed account of the historical development of Lebesgue's theory of measure and integration.

https://doi.org/10.1017/S0266267118000536 Published online by Cambridge University Press

1.2. Issues arising

Let us distinguish two cases depending on the cardinality of S: If S contains only *countably* many elements (e.g. $S = \mathbb{N}$), then it is interesting to note that μ cannot be both CA and uniformly distributed (or, for that matter, μ cannot be a measure that assigns 0 to all singletons). Indeed, let $\{s_1, s_2, \ldots\}$ be an enumeration of all the elements in S. Suppose that μ is uniformly distributed on S and is real-valued, then it must be that $\mu(s_i) = 0$ for all $i \in \mathbb{N}$. But, by CA, $1 = \mu(S) = \mu\left(\bigcup_{i=1}^{\infty} \{s_i\}\right) = \sum_{i=1}^{\infty} \mu(s_i) = 0$, which is absurd. Hence there does not exist a CA uniform distribution on a countable set. It turns out that this simple mathematical fact became one of the main reasons that led de Finetti to reject CA. We shall revisit this line of argument in the context of Savage's subjectivist theory in Section 2.

If, on the other hand, S contains *uncountably* many elements (e.g. $S = \mathbb{R}$), it is known that an extension of Lebesgue measure exists if and only if there exists a measure on the continuum (or any S with $|S| = 2^{\aleph_0}$) satisfying conditions (i)–(iii). Hence, the revised measure problem would be solved if the latter question could be answered. By referring to a result of Ulam (1930) (cf. fn 13 below), Savage gave a definitive answer in saying that such an extension does not exist. This conclusion is inaccurate. In fact, there is no straightforward answer to this question: the existence of a CA measure on $\mathcal{P}(S)$ that extends the Lebesgue measure depends on the background set-theoretic axioms one presupposes. The issue is closely related to the theory of large cardinals; we shall return to this with more details in Section 3.

All in all, these claims of the non-existence of CA measures on $\mathcal{P}(S)$ for both the countable and the uncountable cases lead to the suggestion of weakening the additivity condition (iii) and replacing it with the following condition.

(iii*) μ is finitely additive, that is, for any $X, Y \subseteq S$, if $X \cap Y = \varnothing$ then

$$\mu(X \cup Y) = \mu(X) + \mu(Y). \tag{2}$$

It is plain that (iii) implies (iii*) but not vice versa, hence this condition amounts to placing a weaker constraint on the additivity condition for probability measures. Further, it turns out, *with all other mathematical conditions being equal*,[4] there do exist FA probability measures definable on $\mathcal{P}(S)$ for both the countable and the uncountable cases. These claimed advantages of FA over CA eventually led Savage to opt for FA.

The remainder of the paper is as follows. In Section 2, I review issues concerning uniformly distributed measures within Savage's theory. I observe that it is ill-placed to consider such measures in Savage's system given his own view on uniform partitions. In Section 3, I provide a critical assessment of Savage's set-theoretic arguments in favour of FA, where I defend an account of conceptual realism regarding mathematical practices in applied sciences like decision theory. In Section 4, I explore some technical properties of CA in Savage's system. Section 5 concludes.

[4]This quantifier is important because, as we shall see, whether FA or CA forms a consistent theory depends on other background conditions. Extra mathematical principles are often involved in selecting these conditions when they are in conflict. The discussion on conceptual realism below is an attempt to strike a balance between various mathematical and conceptual considerations.

https://doi.org/10.1017/S0266267118000536 Published online by Cambridge University Press

2. Uniform distributions

De Finetti (1937b) proposed an operationalist account of subjective probability – i.e. the well-known 'betting interpretation' of probability – and showed that a rational decision maker affords the possibility of avoiding exposure to a sure loss if and only if the set of betting quotients with which they handle their bets satisfies conditions (i), (ii) and (iii*) above.

More precisely, let S be a space of possible states, $\mathcal{F}$ be some algebra equipped on S, and members of $\mathcal{F}$ are referred to as events. An event E is said to *occur* if $s \in E$, where S is the true state of the world. Let $\{E_1, \ldots, E_n\}$ be a finite partition of S where each $E_i \in \mathcal{F}$. Further, let $\mu(E_i)$ represent the decision maker's degree of belief in the occurrence of E_i. In de Finetti's theory, an agent's degree of belief (subjective probability) $\mu(E_i)$ is assumed to guide their decisions in betting situations in the following way. $\mu(E_i)$ is the rate at which the agent (the *bettor*) is willing to bet on whether E_i will occur. The bet is so structured that it will cost the bettor $c_i\mu(E_i)$ with the prospect of gaining c_i if event E_i transpires. The c_is are, however, decided by the bettor's opponent (the *bookie*) and can be either positive or negative (a negative gain means that the bettor has to pay the absolute amount $|c_i|$ to the bookie).

The bettor is said to be *coherent* if there is no selection of $\{c_i\}_{i=1}^n$ by the bookie such that $\sup_{s\in S} \sum_{i=1}^n c_i[\chi_{E_i}(s) - \mu(E_i)] < 0$, where χ_{E_i} is the characteristic function of E_i. In other words, the agent's degrees of beliefs in $\{E_i\}_{i=1}^n$ are coherent if no sequence of bets can be arranged by the bookie such that they constantly yield a negative total return for the bettor regardless which state of the world transpires. Guided by this *coherence principle*, de Finetti showed that there exists at least one measure μ defined on $\mathcal{F}$ such that, for any selection of payoffs $\{c_i\}_{i=1}^n$, $\sup_{s\in S} \sum_{i=1}^n c_i[\chi_{E_i}(s) - \mu(E_i)] \geq 0$.

In addition, it was shown by de Finetti (1930) that μ can be extended to any algebra of events that contains $\mathcal{F}$. In particular, μ can be extended to $\mathcal{P}(S)$ – so condition (i) can be satisfied; and that μ is a FA probability measure – that is, μ satisfies (ii) and (iii*). These mathematical results developed by de Finetti in the 1920–30s played an important role in shaping his view on the issue of additivity.[5]

Savage enlists the early works of de Finetti, as well as a similar result proved by Banach (1932), as part of his mathematical reasons not to impose CA. 'It is a little better,' he says, 'not to assume countable additivity as a postulate.' Let us group these main mathematical arguments as follows.

(†) There does not exist a CA uniform distribution over the integers, whereas in the case of FA such a distribution does exist.

(‡) There does not exist, according to Savage, a CA extension of the Lebesgue measure to all subsets of the reals, whereas in the case of FA such an extension does exist.

I shall address (†) in the rest of this section, (‡) the next.

[5]The first chapter of de Finetti (1937b, English translation as de Finetti 1937a) contains a nontechnical summary of de Finetti (1930, 1931, in Italian) on 'the logic of the probable' where the aforementioned mathematical results were given. See Regazzini (2013) for a more detailed account of de Finetti's critique of CA.

https://doi.org/10.1017/S0266267118000536 Published online by Cambridge University Press

2.1. Open-mindedness and symmetry

Appendix A includes an example of a probability measure defined for all subsets of the natural numbers which exhibits the following main properties advocated by de Finetti: (a) it is FA but not CA; (b) it assigns measure 0 to all singletons; and (c) it is uniformly distributed (cf. Example A.1). Opinions, however, differ widely as to why each of these properties constitutes a rational constraint on the notion of subjective probability, especially in view of the operationalist account of probability we owe to writers like Ramsey, Savage and, surely, de Finetti himself.[6] For our purposes, let us focus on uniformity.

De Finetti's insistence on the inclusion of uniform distributions is by and large based on the consideration of *open-mindedness* – rational agents should not have prior prejudices over different types of distributions, all distributions should at least be considered permissible. In addition, the assumption of uniformity is often justified on the ground of certain *symmetry* considerations, that is, in the absence of any existing choice algorithm, each member of a set over which a probability distribution is defined can be seen as having equal possibilities of being chosen. Hence, the argument goes, if there is any incompatibility between uniformity and CA, the latter should yield to the former because of these 'higher' considerations.[7]

Admittedly, as a plea for inclusiveness, it is no doubt quite an attractive idea to be all-embracing – Savage's own demand for μ to be definable for *all* subsets of S is one such example (more discussion on this below), so is the demand for *all* distributions to be considered permissible if not mandatory. This call for openness may resonate positively at first, yet upon a closer examination it is unclear on what grounds this claimed liberalism principally constrains rational decision making. To put it plainly, in making a decision why does anyone have to be subject to the constraints of open-mindedness and symmetry at all? Advocates who appeal to this line of justification for the use of uniform distributions in argument (†) hence face the problem of explaining why a rule for rational action should always respect these 'higher' mandates.

In fact, in the same spirit of liberty, if one is not dogmatic about the criteria of open-mindedness and symmetry – i.e. they are only permissible but not mandatory, to use the same terminologies – then it is easy to see that there is ample room for casting doubts on these principles. Howson and Urbach (1993), for instance, challenge the basis on which a decision maker randomly chooses an integer and treats the choices as being equal: 'any process [of selecting a number]', they say, 'would inevitably be biased towards the 'front end' of the sequence of positive integers'. Indeed, consider, for instance, 1729 and $2^{77,232,917}-1$, both are famous numbers by now.[8] However, before the latter was discovered, it would take a considerable stretch of imagination to envisage a situation where the two numbers are treated as equal citizens of the number kingdom: for one thing, the second is a number with 23, 249, 425 digits, it would take about 6,685 pages of A4 paper to print it

[6]Williamson (1999), for instance, provides a generalized Dutch book argument – an equivalent of the aforementioned coherence principle – for the countable case which naturally leads to CA in a de Finetti-style betting system. Bartha (2004) questions the requirement of assigning measure 0 to all singletons and develops an alternative interpretation for which such a requirement is relaxed.

[7]Thanks to a referee for highlighting this objection.

[8]$2^{77,232,917}-1$ was the largest known prime number when this paper was produced, which was discovered in 2017.

https://doi.org/10.1017/S0266267118000536 Published online by Cambridge University Press

out fully. It is not difficulty to imagine that, before its discovery, this number hardly appeared anywhere on the surface of planet earth, let alone being considered as equally selectable by some choice procedure.

In practice, it has become a matter of taste, so to speak, for theorists to either endorse or reject one or both of these principles, based on their intuitions as well as on what technical details their models require. In fact, Savage is among those who contest the assumption of uniformity. He says:

> [I]f, for example (following de Finetti [D2]), a new postulate asserting that S can be partitioned into an arbitrarily large number of equivalent subsets were assumed, it is pretty clear (and de Finetti explicitly shows in [D2]) that numerical probabilities could be so assigned. It might fairly be objected that such a postulate would be *flagrantly ad hoc*. On the other hand, such a postulate could be made relatively acceptable by observing that it will obtain if, for example, in all the world there is a coin that the person is firmly convinced is fair, that is, a coin such that any finite sequence of heads and tails is for him no more probable than any other sequence of the same length; though such coin is, to be sure, *a considerable idealization*. (p. 33, emphases added. [D2] refers to de Finetti 1937b)

As seen, for Savage, only a thin line separates the assumption of uniformity from a spell of good faith. There is yet another and more technical reason why he refrains from making this assumption as it is alluded to in the quote above. This will take a little unpacking.

In deriving probability measures, both de Finetti and Savage invoke a notion of *qualitative probability*, which is a binary relation $\succeq$ defined over events: For events E, F, $E \succeq F$ says that 'E is weakly more probable than F' (or 'E is no less probable than F'). E and F are said to be equally probable, written $\mathrm{E} \equiv \mathrm{F}$, if both $E \succeq F$ and $F \succeq E$. A qualitative probability satisfies a set of intuitive properties.[9] The goal is to impose some additional assumption(s) on $\succeq$ so that it can be represented by a unique *quantitative probability*, μ, that is, $E \succeq F$ if and only if $\mu(E) \geq \mu(F)$.

The approach adopted by de Finetti (1937b) and Koopman (1940a, 1940b, 1941) was to assume that any event can be partitioned into arbitrarily small and equally probable sub-events – i.e. the assumption of *uniform partitions* (UP):

> **UP:** For any event A and any $n < \infty$, A can be partitioned into n many equally probable sub-events, i.e. there is a partition $\{B_1, \ldots, B_n\}$ of A such that $B_i \equiv B_j$.

As noted by Savage, when added to the list of properties of qualitative probabilities, UP is deductively sufficient in delivering a numeric representation (UP is a simple

[9]Formally, a binary relation $\succeq$ over an algebra of events $\mathcal{F}$ is said to be a *qualitative probability* if the following hold for all $A, B, C \in \mathcal{F}$: (1) $\succeq$ is a total preorder; (2) $A \succeq \varnothing$; (3) $S \succ \varnothing$; and (4) if $A \cap C = B \cap C = \varnothing$ then $A \succeq B \Longleftrightarrow A \cup C \succeq B \cup C$.

Historically, it was thought that, at least for the finite cases, these four conditions are sufficient in arriving at numeric representations. A counterexample, however, was quickly found by Kraft *et al.* (1959) who gave an example of a qualitative probability defined over the Boolean algebra of all subsets of a set consisting of only five members, for which there is no numeric representation.

https://doi.org/10.1017/S0266267118000536 Published online by Cambridge University Press

version of the *Archimedean* condition normally required in a representation theorem). Thus, it is out of *both* the intuitive appeal to symmetry *and* mathematical necessity that de Finetti comes to endorse UP.

Savage, however, is not tied to either of these two considerations. Given his view on uniform distributions – they are being 'flagrantly ad hoc' – Savage needs to find an alternative way of arriving at a representation theorem without invoking UP. To this end, he introduces a concept of *almost uniform partition* (AUP):

> **AUP:** For any event A and any $n < \infty$, A can be partitioned into n many sub-events such that the union of any $r < n$ sub-events of the partition is no more probable than that of any $r + 1$ sub-events.

The idea is to partition an event into arbitrarily 'fine' sub-events but without asking them to be equally probable. It is plain that a uniform partition is an almost uniform partition, but not vice versa.

The genius of Savage's proof was thus to show (1) that his proposed axioms are sufficient in defining AUP in his system and (2) that AUP is all that is needed in order to arrive at a numeric representation of qualitative probability. Technical details of how numeric probability measures are derived in Savage's theory do not concern us here.[10] But what is clear is that, in view of Savage's take on uniformity it would be misplaced to invoke argument (†) against CA within his system.

2.2. Money pump and infinite bets

Even if we grant that uniformly distributed measures be permissible in Savage-type decision models, it can be shown that the admission of such a measure together with FA may subject an agent to a Dutch book. Appendix B contains an example of Adams' money pump, where the agent's subjective probability is FA but not CA and is defined in terms of the uniform distribution λ introduced in Example A.1. As shown there, a series of bets can be arranged such that the bettor suffers a sure loss.

Adams' money pump is surprising, because it results in a set of incoherent choices made by the bettor, a result that is precisely what subjectivist systems like Savage's are devised to avoid. Indeed, if there is a single principle that unites Bayesian subjectivists it would arguably be *coherence*. One important reason this principle is fundamental to Bayesian subjectivism is that this coherence-oriented approach to rationality democratizes, so to speak, legitimate reasons for acting: different agents might be moved by different (probabilistic or utility) considerations that lead them to behave differently (or similarly), but as long as their behaviours are consistent, they are equally rational in the subjectivist sense. For instance, when both picking up a slice of cheesecake, I doubt my five-year-old nephew had to endure the same kind of inner struggles that I had to deal with – he, for one thing, certainly would not think of going for the second largest slice for a quite nuanced reason. But as long as we both are consistent in our choices we are equally rational in

[10]For an outline see see Gaifman and Liu (2018: §3).

https://doi.org/10.1017/S0266267118000536 Published online by Cambridge University Press

the eyes of subjectivists. For subjectivists, the notion of rationality is based on a rather weak logical principle, namely coherence.[11] Adams' example reveals a conflict with this basic principle.

When faced with this difficulty, advocates of FA often argue that given that a subjective decision theory is a systematization of coherent decision making by rational agents it is unclear what it means for a bettor to fulfil the task of coherently betting *infinitely* many times. On this view, the challenge from Adams' money pump, which requires the bettor to accept infinitely many gambles, is a non-starter, for it envisages a situation that is not operationally feasible.[12]

Perhaps, this is another place where a theorist needs to exercise their acquired taste in order to discern whether or not it is conceptually salient to entertain infinite bets. I, nonetheless, want to point out that there is already a great deal of idealization built into the subjectivist theory that requires us to be somewhat imaginative. As a starter, it is a common practice to assume that the agents being modelled are *logically omniscient*. One reason for upholding such an assumption is that, in a model of decision making under uncertainties, uncertainties may stem from different sources: they may be due to the lack of empirical knowledge or the result of computational failures. To separate different types of uncertainties, it is often assumed, like in de Finetti's and Savage's systems, that the agents are equipped with unlimited deductive capacities in logic.

However, *it would be quite a double standard* to insist in these systems that agents are infinitely capable when it comes to computational or inferential performances, on the one hand; but to appeal to certain physical or psychological realism when it comes to accepting bets, on the other. This, nevertheless, does not mean that there can be no limits to the amount of idealizations that one injects into a model. In what follows we will explore some measures for constraining idealism that may lead to a better understanding of theoretic modelling.

3. Higher mathematics and conceptual realism

Let us return to argument (‡). Savage (1972: 41) cites the well-known result of Ulam (1930) in asserting that any atomless σ-additive extension of Lebesgue measure to *all* subsets of the unit interval is incompatible with the continuum hypothesis (CH), from which he concludes that there is no extension that satisfies *all* of (i)–(iii). However, it is unclear why this constitutes a sufficient reason for relaxing CA.[13]

[11]This, of course, does not mean that my rational decisions – rational in the subjectivist sense – are always wise ones. I might get totally soaked in the rain due to some bad, albeit coherent, estimations I made earlier. But in hearing my complaints, the subjectivist would simply shrug and point out that they never promised to keep my pants dry.

[12]Thanks to Isaac Levi for pointing out this line of objection to me, which echoes de Finetti's view that a rational agent is not obliged to accept more than finitely many fair bets at a given time.

[13]Ulam (1930) proved that, for any uncountable set S with $|S| = \kappa$, it can be shown in ZFC that if κ is a successor (and hence a regular) cardinal (e.g. $\kappa = \aleph_1$), then there does not exist a measure on S satisfying all of (i)–(iii). It follows that if there is a σ-additive non-trivial extension of the Lebesgue measure on $2^{\aleph_0}$ then CH must fail. (It is worth mentioning that, prior to Ulam, Banach and Kuratowski (1929) showed that if there is a measure on $2^{\aleph_0}$ then $2^{\aleph_0} > \aleph_1$.) Yet, even without the concern for CH, there is an aspect of Ulam's original results that was not addressed by Savage: it was shown by Ulam (1930) that, in ZFC,

https://doi.org/10.1017/S0266267118000536 Published online by Cambridge University Press

Table 1. Consistent extensions of the Lebesgue measure to $\mathcal{P}(\mathbb{R})$ under different set-theoretic assumptions

Axioms	Savage	Solovay
FA	✓	
CA		✓
AC	✓	
DC		✓
CH	✓	
I		✓
$\mathcal{P}(\mathbb{R})$	✓	✓

As a matter of fact, in his article entitled 'A model of set theory in which every set of reals is Lebesgue measurable', Solovay (1970) showed that such a σ-additive extension of the Lebesgue measure to all sets of reals *does* exist if the existence of an inaccessible cardinal (I) and a weaker version of AC – i.e. the principle of dependent choice (DC) – are assumed.[14] Thus, it seems that insofar as the possibility of obtaining a σ-additive extension of the Lebesgue measure to all subsets of the reals is concerned, Savage's set-theoretic argument – which calls for exclusion of CA – is inconclusive. For the existence of such an extension really depends on the background set theory: it does not exist in ZFC + CH, but does exist in ZF + DC (assuming ZFC + I is consistent) (cf. Table 1 for a side-by-side comparison).

3.1. Logical omniscience and measurability

As seen, Savage's set-theoretic argument for not imposing CA was given in ZFC + CH, where it is known that, in the uncountable case, there is no non-trivial measure that simultaneously satisfies conditions (i)–(iii) above; and that Savage's immediate reaction was to replace the third, i.e. the CA condition, with FA.[15]

if there is a σ-additive non-trivial measure μ on any uncountable set S with $|S| = \kappa$ then μ is a measure on κ such that

(1) either κ is a measurable cardinal (and hence an inaccessible cardinal), on which a non-trivial σ-additive two-valued measure can be defined;
(2) or κ is a real-valued measurable cardinal (and hence a weakly inaccessible cardinal) such that $\kappa \leq 2^{\aleph_0}$, on which a non-trivial σ-additive atomless measure can be defined.

In the second case, it is plain that μ can be extended to a measure on $2^{\aleph_0}$: for any $X \subseteq 2^{\aleph_0}$, let $\mu(X) = \mu\ (X \cap \kappa)$. This leads to a general method of obtaining a countably additive measure on all subsets of the reals that extends the Lebesgue measure (see Jech 2003: 131).

[14]The relative consistency proof by Solovay (1970) showed that if ZFC + I has a model then ZF + DC has a model in which every set of reals is Lebesgue measurable (see also Jech 2003: 50).

[15]Savage was not alone in taking this line. Seidenfeld (2001), for instance, enlisted the non-existence of a non-trivial σ-additive measure defined over the power set of an uncountable set as the first of his six reasons for considering a theory of probability that is based on FA. It is interesting to note that Seidenfeld also referred to the result of Solovay, however no further discussion on the significance of this result on CA

https://doi.org/10.1017/S0266267118000536 Published online by Cambridge University Press

The particular set-theoretic argument Savage relied on – namely the existence of the Ulam matrix which leads to the non-existence of a measure over $\aleph_1$ (cf. fn 13) – uses AC in an essential way.

One unavoidable consequence of this approach is that it introduces non-measurable sets in ZFC. Now, if one insists on imposing condition (i) for defining subjective probability – that is, a subjective probability measure be defined for *all* subsets of the states – this amounts to introducing non-measurable sets into Savage's decision model. Yet, it is unclear what one may gain from making such a high demand.

Non-measurable sets are meaningful only insofar as we have a good understanding of the contexts in which they apply. These sets are highly interesting within certain branches of mathematics largely because the introduction of these sets reveals in a deep way the complex structures of the underlying mathematical systems. However, this does not mean that these peculiar set-theoretic constructs should be carried over to a system that is primarily devised to model rational decision making.

It might be objected that the subjectivist theory we are concerned with is, after all, highly idealized – the assumption of logical omniscience, for instance, is adopted in almost all classical decision models. By that extension, one may as well assume that the agent being modelled is *mathematically omniscient* (an idealized perfect mathematician). Then, it should be within the purview of this super agent to contemplate non-measurable sets/events in decision situations. Besides, if we exclude non-measurable sets from decision theory, then why stop there? In other words, precisely where shall we draw the line between what can and cannot be idealized in such models?

This is a welcome point, for it highlights the issue of how much idealism one can instil into a theoretic model. It is no secret that decision theorists are constantly torn between the desire to aspire to the perfect rationality prescribed in the classical theories, on the one hand; and the need for a more realistic approach to practical rationality, on the other. Admittedly, this 'line' between idealism and realism is sometimes a moving target – it often depends on the goal of modelling and, well, the fine taste of the theorist. An appropriate amount of idealism allows us to simplify the configuration of some aspects of a theoretic model so that we can focus on some other aspects of interest. An overdose, however, might be counter-effective as it may introduce unnecessary mysteries into the underlying system.

Elsewhere (Gaifman and Liu 2018), we introduced a concept of *conceptual realism* in an attempt to start setting up some boundaries between idealism and realism in decision theory. We argued that, as a minimum requirement, the idealizations one entertains in a theoretic model should be justifiable within the confines of one's conceivability. And, importantly, what is taken to be conceivable should be anchored from the theorist's point of view, instead of delegating it to some imaginary super agent. Let me elaborate with the example at hand.

As noted above, it is a common practice to adopt the assumption of logical omniscience in Bayesian models. This step of idealization, I stress, is not a blind leap of faith. Rather, it is grounded in our understanding of the basic logical and computational

versus FA was given (see Seidenfeld 2001: 168). See Bingham (2010: §9) for a discussion and responses to each of Seidenfeld's six reasons. The set-theoretic argument presented here, in response to Seidenfeld's first, i.e. the measurability reason is different from Bingham's 'approximation' argument.

https://doi.org/10.1017/S0266267118000536 Published online by Cambridge University Press

apparatuses involved. We, as actual reasoners, acknowledge that our inferential performances are bounded by various physical and psychological limitations. Yet a good grasp of the underlying logical machinery gives rise to the conceivable picture as to what it means for a logically omniscient agent to fulfil, at least *in principle*, the task of, say, drawing all the logical consequences from a given proposition.[16]

This justificatory picture – which apparently is based on certain *inductive reasoning* – becomes increasingly blurry when we start contemplating how our super agent comes to entertain non-measurable events in the context of rational decision making. The Banach–Tarski paradox sets just the example of how much we lack geometric and physical intuitions when it comes to non-measurable sets. This means, unlike logical omniscience, we don't have any clue of what we are asking from our super agent: beyond any specific set-theoretic context there is just no good intuitive basis for conceiving non-measurable sets. Yes, it might be comforting that we can shift the burden of conceiving the inconceivables to an imaginary super agent, but in the end such a delegated job is either meaningless or purely fictional. So it seems that if there is any set-theoretic oddity to be avoided here it should be non-measurable sets.[17]

On this matter, I should add that Savage himself is fully aware that the set-theoretic context in which his decision model is developed exceeds what one can expect from a rational decision maker (Savage 1967). He also cites the Banach–Tarski paradox as an example to show the extent to which highly abstract set theory can contradict common sense intuitions. However, it seems that Savage's readiness to embrace all-inclusiveness of defining the background algebra as 'all sets of states' overwhelms his willingness to avoid this set-theoretic oddity.

3.2. The standard approach

In fact, the situation can be largely simplified if we choose to work, instead of with *all* subsets of the state space S, but with a sufficiently rich collection of subsets of S (for instance, the Borel sets $\mathscr{B}$ in the case where $S = \mathbb{R}$) where, as a well established theory, CA is in perfectly good health. That is, instead of (i), we require that

(i*) μ is defined on (Lebesgue) measurable sets of $\mathbb{R}$.

[16]With this being said, I should however point out that the demand to have a deductively closed system remains as a challenge to any normative theory of beliefs. In his essay titled 'difficulties in the theory of personal probability', Savage (1967: 308) remarked that the postulates of his theory imply that agents should behave in accordance with the logical implications of all that they know, which can be very costly. In other words, conducting logical deductions is an extremely resource consuming activity – the merits it brings can sometimes be offset by its high costs. Hence some might say that the assumption of logical omniscience – a promise of being able to discharge unlimited deductive resources – may at best be seen as an unfeasible idealization. Nonetheless, this is not the place to open a new line of discussion on the legitimacy of logical omniscience. The point I am trying to make is rather that, unlike non-measurable sets, being logically all powerful, however unfeasible, is not something that is conceptually inconceivable.

[17]In an unpublished work, Haim Gaifman made a similar point against the often cited analogy between finitely additive probabilities in countable partitions and countably additive probabilities in uncountable partitions in the literature (see e.g. Schervish and Seidenfeld 1999) that such an analogy plays a major heuristic role in set theory but provides no useful guideline in the case of subjective probabilities for the reason that certain mathematical structures required to make salient this analogy have no meaning in a personal decision theory, and hence the analogy essentially fails.

https://doi.org/10.1017/S0266267118000536 Published online by Cambridge University Press

Note that the price of forfeiting the demand from condition (i) is a rather small one to pay. It amounts to disregarding all those events that are defined by Lebesgue non-measurable sets. Indeed, even Savage himself conceded that

> All that has been, or is to be, formally deduced in this book concerning preferences among sets, could be modified, *mutatis mutandis*, so that the class of events would not be the class of all subsets of S, but rather a Borel field, that is, a σ-algebra, on S; the set of all consequences would be measurable space, that is, a set with a particular σ-algebra singled out; and an act would be a measurable function from the measurable space of events to the measurable space of consequences. (Savage 1972: 42)

It shall be emphasized that this modification of the definition of events from, say, the set of all subsets of (0,1] to the Borel set $\mathscr{B}$ of (0,1] is not carried out at the expense of dismissing a large collection of events that are otherwise representable. As noted by Billingsley (2012: 23), '[i]n fact, $\mathscr{B}$ contains all the subsets of (0,1] actually encountered in ordinary analysis and probability. It is large enough for all 'practical' purposes.' By 'practical purposes' I take it to mean that all events and measurable functions considered in economic theories in particular are definable using only measurable sets, and, consequently, there is no need to appeal to non-measurable sets.

To summarize, I have shown that Savage's mathematical argument against CA is inconclusive – its validity depends on the background set-theoretic set-ups – and that CA *can* in fact form a consistent theory under appropriate arrangements. This brought us to a fine point concerning how one understands background mathematical details and how best to incorporate idealism into a theoretic model. As we have seen, a healthy amount of conceptual realism may keep us from many highly idealized fantasies.

4. Countably additive probability measure

In this section I will discuss how CA can be introduced into Savage's theory and its relation to other postulates in the system. I will also point out an interesting line for future research. The discussion in this section will assume the standard approach from Section 3.2. Readers who are not interested in this technical part of Savage's theory may safely 'jump to the conclusion' in the last section.

4.1. Quantitative and qualitative continuities

In Savage's representation theorem, it is assumed that the background algebra is a σ-algebra (i.e. closed under infinite unions and intersections). Savage remarks that

> It may seem peculiar to insist on σ-algebras as opposed to finitely additive algebras even in a context where finitely additive measures are the central objects, but countable union do seem to be essential to some theorems of §3 ... (Savage 1972: 43)

https://doi.org/10.1017/S0266267118000536 Published online by Cambridge University Press

This 'peculiar' feature of the system reveals a certain 'imbalance' between the finite adding speed, so to speak, of FA, on the one hand; and countable unions in a σ-algebra, on the other. There are two intuitive ways to restore the balance. One is to reduce the size of the background algebra. This is the approach adopted in Gaifman and Liu (2018), where we invented a new technique of 'tri-partition trees' and used it to prove a representation theorem in Savage's system without requiring the background algebra to be a σ-algebra.

The other option is to tune up the adding speed and bring 'continuity' between probability measures and algebraic operations. To be more precise, let $(S, \mathcal{F}, \mu)$ be a measure space, μ is said to be *continuous from below*, if, for any sequence of events $\{A_n\}_{n=1}^{\infty}$ and event A in $\mathcal{F}$, $A_n \uparrow A$ implies that $\mu(A_n) \uparrow \mu(A)$; it is *continuous from above* if $A_n \downarrow A$ implies $\mu(A_n) \downarrow \mu(A)$, and it is *continuous* if it is continuous from both above and below.[18] It can be shown that continuity fails in general, if μ is merely FA.[19] In fact, it can be proved that continuity holds if and only if μ is CA.

One way to introduce continuity to Savage's system is to impose a stronger condition on qualitative probability. Let $\succeq$ be a qualitative probability, defined on a σ-algebra $\mathcal{F}$ of the state space S. Following Villegas (1964), $\succeq$ is said to be *monotonously continuous* if, given any sequence of events $A_n \uparrow A$ ($A_n \downarrow A$) and any event B,

$$A_n \preceq B (A_n \succeq B) \text{ for all } n \Longrightarrow A \preceq B (A \succeq B). \tag{3}$$

Moreover, Villegas showed that if a qualitative probability $\succeq$ is atomless and monotonously continuous then the numerical probability μ that agrees with $\succeq$ is unique and CA.[20] Since the qualitative probability in Savage's system is non-atomic, it is sufficient to introduce the the property of monotone continuity in order to bring in CA. We thus can add the following postulate, P8, to Savage's P1–P7 (in Appendix C), which is a reformulation of (3) in terms of preferences among Savage acts.

P8: For any $a,b \in X$ and for any event B and any sequence of events $\{A_n\}$,

(i) if $A_n \uparrow A$ and $\mathfrak{c}_a|A_n + \mathfrak{c}_b|\overline{A_n} \preccurlyeq \mathfrak{c}_a|B + \mathfrak{c}_b|\overline{B}$ for all n then $\mathfrak{c}_a|A + \mathfrak{c}_b|\overline{A} \preccurlyeq \mathfrak{c}_a|B + \mathfrak{c}_b|\overline{B}$;
(ii) if $A_n \downarrow A$ and $\mathfrak{c}_a|A_n + \mathfrak{c}_b|\overline{A_n} \succcurlyeq \mathfrak{c}_a|B + \mathfrak{c}_b|\overline{B}$ for all n then $\mathfrak{c}_a|A + \mathfrak{c}_b|\overline{A} \succcurlyeq \mathfrak{c}_a|B + \mathfrak{c}_b|\overline{B}$.

This allows us to introduce CA into Savage's decision model as an added postulate. P8 (and hence CA), however, cannot replace the role played by P7. This is a delicate matter, we investigate it in the next section.

[18]As notational conventions, $A_n \uparrow A$ means that $A_1 \subseteq A_2 \subseteq \cdots$ and $\bigcup_i A_i = A$, and $\mu(A_n) \uparrow \mu(A)$ means that $\mu(A_1) \leq \mu(A_2) \leq \cdots$ and $\mu(A_n) \to \mu(A)$ as $n \to \infty$. Similarly, for the other case.

[19]As an illustration, it is interesting to note that the strictly FA measure λ in Appendix A is neither continuous from above nor from below.

[20]Villegas (1964) showed that monotone continuity is a necessary and sufficient condition for the agreeing numerical measure to be CA. It was further shown that any qualitative probability defined on a finite algebra can be extended to a qualitative probability on a σ-algebra satisfying monotone continuity, fineness, tightness. Thanks to Savage's P6, the qualitative probabilities derived in the system are atomless, fine and tight, hence CA obtains if the monotone continuity is in place.

https://doi.org/10.1017/S0266267118000536 Published online by Cambridge University Press

4.2. Countable additivity and P7

In Savage's theory, a representation theorem for simple acts (i.e. acts that may potentially lead to only finitely many consequences) can be given under P1–P6. Savage's P7 plays the sole role of extending utility representations from simple acts to general acts (i.e. acts that may potentially lead to infinitely many different outcomes).

Savage (1972: 78) gave an example which satisfies the first six of his seven postulates but not the last one. This is intended to show that the seventh postulate (P7) is independent of other postulates in Savage's original system. Upon showing the independence of P7, Savage remarked that '[f]inite, as opposed to countable, additivity seems to be essential to this example', and he conjectured that 'perhaps, if the theory were worked out in a countably additive spirit from the start, little or no counterparts of P7 would be necessary'. This section is aimed at taking a closer look at Savage's remark on the relation between CA and utility extension under various versions of P7.

In a footnote to the remark above Savage adds: 'Fishburn (1970, Exercise 21, p. 213) has suggested an appropriate weakening of P7'. It turned out that this is inaccurate. To wit, the following is Fishburn's suggestion (expressed using our notation).

> **P7b:** For any event $E \in \mathcal{F}$ and $a \in X$, if $\mathfrak{c}_a \succcurlyeq_E \mathfrak{c}_{g(s)}$ for all $s \in E$ then $\mathfrak{c}_a \succcurlyeq_E g$; and if $\mathfrak{c}_a \preccurlyeq_E \mathfrak{c}_{g(s)}$ for all $s \in E$ then $\mathfrak{c}_a \preccurlyeq_E g$.

P7b is weaker than P7 in that it compares act g with a constant act instead of another general act f. Note that Fishburn's P7b is derived from the following condition A4b that appeared in his discussion on preferential axioms and bounded utilities (§10.4).

> **A4b:** Let X be a set of prizes/consequences and $\Delta(X)$ be the set of all probability measures defined on X, then for any $P \in \Delta(X)$ and any $A \subseteq X$ if $P(A) = 1$ and, for all $x \in A$, $\delta_x \succcurlyeq (\preccurlyeq) \delta_y$ for some $y \in X$ then $P \succcurlyeq (\preccurlyeq) \delta_y$, where δ_x denotes the probability that degenerates at x.

A4b, together with other preference axioms discussed in the same section, are used to illustrate, among other things, the differences between measures that are countably additive and those are not. It was proved by Fishburn that the *expected utility hypothesis* holds under A4b, that is, if $\Delta(X)$ contains *only* CA measure, then

$$P \succ Q \Longleftrightarrow E(u, P) > E(u, Q), \quad \text{for all } P, Q \in \Delta(X). \tag{4}$$

Fishburn then showed, by way of a counterexample, that the hypothesis fails if the set of probability measures contains also merely FA ones. Because of its direct relevancy to our discussion on the additivity condition below, we reproduce this example (Fishburn 1970: Theorem 10.2) here.

Example 4.1. Let $X = \mathbb{N}^+$ with $u(x) = x/(1 + x)$ for all $x \in X$. Let $\Delta(X)$ be the set of all probability measures on the set of all subsets of X and defined u on $\Delta(X)$ by

$$u(P) = E(u, P) + \inf \left\{ P[u(x) \geq 1 - \epsilon] : 0 < \epsilon \leq 1 \right\}. \tag{5}$$

https://doi.org/10.1017/S0266267118000536 Published online by Cambridge University Press

Define $\succ$ on $\Delta(X)$ by $P \succ Q$ iff $u(P) > u(Q)$. It is easy to show that A4b holds under this definition. However if one takes P to be the measure in Appendix A, i.e. a finitely but not countably additivity probability measure, then we have $u(\lambda) = 1 + 1 = 2$. Hence $u(\lambda) \neq E(u,\lambda) = 1$. This shows the expected utility hypothesis fails under this example. ◁

However, as pointed out by Seidenfeld and Schervish (1983: Appendix), Fishburn's proof of (4) using A4b was given under the assumption that $\Delta(X)$ is closed under countable convex combination (condition S4 in Fishburn 1970: 137), which in fact is not derivable in Savage's system. They show through the following example (see Example 2.3, p. 404) that the expected hypothesis fails under the weakened P7b (together with P1–P6) and this is so even when the underlying probability is CA.

Example 4.2. Let S be [0,1) and X be the set of rational numbers in [0,1). Let μ be uniform probability on measurable subsets of S and let all measurable function f from S to X satisfying $V[f] = \lim_{i \to \infty} \mu[f(s) \geq 1 - 2^{-i}]$ be acts. For any act f, let $U[f] = \int_S u(f)d\lambda$ where $u(x) = x$ is a utility function on X and define

$$W[f] = U[f] + V[f] = \int_S u(f(s))d\mu(s) + \lim_{i \to \infty} \mu\left[f(s) \geq 1 - \frac{1}{2^i}\right]. \quad (6)$$

Further, define $f \succ g$ if $W[f] > W[g]$. It is easy to see that P1–P6 are satisfied. To see that W satisfies P7b, note that if for any event E and any $a \in X$, $\mathfrak{c}_a \succcurlyeq_E \mathfrak{c}_{g(s)}$ for all $s \in E$, then by (6), we have $1 > u(a) \geq u(g(s))$ for any $s \in E$. Note that $1 > u(g(s))$ implies $V[g\chi_E] = 0$ where χ is the indicator function. Thus, $W[\mathfrak{c}_a\chi_E] = \int_E u(a)d\mu \geq \int_E u(g(s))d\mu(s) = W[g\chi_E]$. The case $\mathfrak{c}_a \preccurlyeq_E \mathfrak{c}_{g(s)}$ can be similarly shown. ◁

In other words, contrary to what Savage had thought, P7b is in fact insufficient in bringing about a full utility representation theorem even in the presence of CA.[21] This shows, *a fortiori,* that CA alone is insufficient in carrying the utility function derived from P1–P6 from simple acts to general acts. Seidenfeld and Schervish (1983: Example 2.2) also showed that this remains the case even if the set of probabilities measure is taken to be closed under countable convex combination.

As seen, Savage's last postulate which plays the role of extending the utilities from simple acts to general acts cannot be easily weakened even in the presence of CA. Yet, on the other hand, it is clear that CA is a stronger condition than FA originally adopted in Savage's theory. So, for future work, it might be of interest to find an appropriate weakening of P7 in a Savage-style system in the presence of CA that is still sufficient in extending utilities from simple acts to general acts.

5. Conclusion

Two general concerns often underlie the disagreements over FA versus CA in Bayesian models. First, there is a mathematical consideration as to whether or not the kind of additivity in use accords well with demanded mathematical details in the background. Second, there is a philosophical interest in understanding whether or not the type of additivity in use is conceptually justifiable.

[21]Here I am indebted to Teddy Seidenfeld.

https://doi.org/10.1017/S0266267118000536 Published online by Cambridge University Press

Regarding the first concern, Savage provided a set-theoretic argument against CA where he argues that CA is not in good conformity with demanded set-theoretic details. As we have seen, Savage's argument is misguided due to an overlook of some crucial technical details: CA *can* be coherently incorporated in his personal decision theory without serious set-theoretic complications. As for the second concern, we noted that both arguments (†) and (‡) – which had been enlisted to bear evidence against CA in Savage-type decision theory – are inadequate in the context of Savage's theory of expected utilities. In dealing with these issues, we took a closer look at the mathematical details involved, where I argued that in order for a piece of mathematics employed in defining subjective probability to be meaningful it is necessary that it be handled in a conceptually realistic manner anchored from the theorist's point of view.

As far as Savage's system is concerned, there does not seem to be sufficient reason why the model cannot be extended to include CA. As a general guide, it might be of interest to take a more pragmatic line when it comes to adopting FA or CA. Following this advice, I would return to a point made at the outset that, given how widespread and useful CA measures are in modern probability theory, it's better to presuppose CA and only to weaken it to FA when called for.

Author ORCID. Yang Liu 0000-0001-8865-4647

Acknowledgements. I am grateful to Jean Baccelli, Haim Gaifman, Issac Levi, Michael Nielsen, Teddy Seidenfeld and Rush Stewart for helpful discussions on early versions of this paper; and to two anonymous reviewers for many astute comments. The research of this paper was partially supported through a grant from the Templeton World Charity Foundation (TWCF0128) and by the Leverhulme Trust.

References

Adams E. W. 1962. On rational betting systems. *Archive for Mathematical Logic* **6**, 7–29.

Banach S. 1932. *Théorie des opérations linéaires*. Warszawa: Z subwencji funduszu kultury narodowej.

Banach S. and K. Kuratowski 1929. Sur une généralisation du problème de la mesure. *Fundamenta Mathematicae* **14**, 127–131.

Bartha P. 2004. Countable additivity and the de Finetti lottery. *British Journal for the Philosophy of Science* **55**, 301–321.

Billingsley P. 2012. *Probability and Measure*. Anniversary Edition. Chichester: Wiley.

Bingham N. H. 2000. Studies in the history of probability and statistics XLVI. Measure into probability: from Lebesgue to Kolmogorov. *Biometrika* **87**, 145–156.

Bingham N. H. 2010. Finite additivity versus countable additivity. *Electronic Journal for History of Probability and Statistics* **6**, 1–35.

de Finetti B. 1930. Problemi determinati e indeterminati nel calcolo delle probabilità. *Atti R. Accademia Nazionale dei Lincei, Serie VI, Rend* **12**, 367–373.

de Finetti B. 1931. Sul significato soggettivo della probabilità. *Fundamenta Mathematicae* **17**, 298–329.

de Finetti B. 1937a. Foresight: its logical laws, its subjective sources. In H. E. Kyburg and H. E. Smokler (eds), *Studies in Subjective Probability*, 2nd Edn. Huntington, NY: Robert E. Krieger Publishing Co., pp. 53–118. 1980.

de Finetti B. 1937b. La prévision: ses lois logiques, ses sources subjectives. *Annales de l'Institut Henri Poincaré* **7**, 1–68.

Fishburn P. C. 1970. *Utility Theory for Decision Making*. New York, NY: Wiley.

Gaifman H. and Y. Liu 2018. A simpler and more realistic subjective decision theory. *Synthese* **195**, 4205–4241.

Hawkins T. 1975. *Lebesgue's Theory of Integration: Its Origins and Development.* , 2nd Edn. New York, NY: Chelsea Pub. Co.

Howson C. and P. Urbach 1993. *Scientific Reasoning: The Bayesian Approach*. La Salle, IL: Open Court.

https://doi.org/10.1017/S0266267118000536 Published online by Cambridge University Press

Hrbacek K. and T. J. Jech 1999. *Introduction to Set Theory. Chapman & Hall/CRC Pure and Applied Mathematics*, 3rd Edn. New York, NY: Marcel Dekker.

Jech T. J. 2003. *Set Theory: The Third Millennium Edition, Revised and Expanded*. Berlin: Springer-Verlag.

Koopman B. O. 1940a. The axioms and algebra of intuitive probability. *Annals of Mathematics* **41**, 269–292.

Koopman B. O. 1940b. The bases of probability. *Bulletin of the American Mathematical Society* **46**, 763–774.

Koopman B. O. 1941. Intuitive probabilities and sequences. *Annals of Mathematics* **42**, 169–187.

Kraft C., J. Pratt and A. Seidenberg 1959. Intuitive probability on finite sets. *Annals of Mathematical Statistics* **30**, 408–419.

Lebesgue H. 1902. Intégrale, longueur, aire. *Annali di Matematica* **3**, 231–359.

Rao K. P. S. B. and M. B. Rao 1983. *Theory of Charges: A Study of Finitely Additive Measures*. New York, NY: Academic Press.

Regazzini E. 2013. The origins of de Finetti's critique of countable additivity. *Advances in Modern Statistical Theory and Applications: A Festschrift in Honor of Morris L. Eaton* **10**, 63–82.

Savage L. J. 1954. *The Foundations of Statistics*. New York, NY: John Wiley & Sons.

Savage L. J. 1967. Difficulties in the theory of personal probability. *Philosophy of Science* **34**, 305–310.

Savage L. J. 1972. *The Foundations of Statistics*, 2nd revised edn. New York, NY: Dover Publications.

Schervish M. J. and T. Seidenfeld 1999. Statistical implications of finitely additive probability. In J. B. Kadane, M. J. Schervish and T. Seidenfeld (eds), *Rethinking the Foundations of Statistics*. Cambridge: Cambridge University Press, pp. 211–232.

Seidenfeld T. 2001. Remarks on the theory of conditional probability: some issues of finite versus countable additivity. In V. F. Hendricks, S. A. Pedersen and K. F. Jorgensen (eds), *Probability Theory: Philosophy, Recent History and Relations to Science*. Dordrecht: Kluwer, pp. 167–178.

Seidenfeld T. and M. J. Schervish 1983. A conflict between finite additivity and avoiding Dutch book. *Philosophy of Science* **50**, 398–412.

Solovay R. M. 1970. A model of set-theory in which every set of reals is Lebesgue measurable. *Annals of Mathematics* **92**, 1–56.

Ulam S. 1930. Zur masstheorie in der allgemeinen mengenlehre. *Fundamenta Mathematicae* **16**, 140–150.

Villegas C. 1964. On qualitative probability *s*-algebras. *Annals of Mathematical Statistics* **35**, 1787–1796.

Vitali G. 1905. *Sul problema della misura dei gruppi de punti di una retta*. Bologna: Tipografia Gamberini e Parmeggiani.

Williamson J. 1999. Countable additivity and subjective probability. *British Journal for the Philosophy of Science* **50**, 401–416.

Appendix A. Uniform distribution over the natural numbers

A uniformly distributed probabilistic measure on natural numbers $\mathbb{N}$ is of particular interest because (1) it serves a good purpose of delineating the difference between finite additivity and countable additivity; (2) its use is often tied to the notion of randomness: it amounts to saying choose a number 'at random'. The latter is commonly understood in the following relative frequentist interpretation of uniformity of natural numbers. Let A be any subset of $\mathbb{N}$. For each number $n < \infty$, denote the number of elements in A that are less or equal to n by $A(n)$, that is,

$$A(n) = |A \cap \{1, \ldots, n\}|. \tag{A.1}$$

Define the *density* of A by the limit (if exists)

$$d(A) = \lim_{n \to \infty} \frac{A(n)}{n}. \tag{A.2}$$

Let $\mathcal{C}_d$ be the collection of all sets of natural numbers that have densities. The following properties of the density function are easy to verify.

(1) $d(\varnothing) = 0$ and $d(\mathbb{N}) = 1$.
(2) For each natural number n, $d(\{n\}) = 0$.
(3) For any finite $A \in \mathcal{C}_d$, $d(A) = 0$.
(4) If $A, B, A \cup B \in \mathcal{C}_d$ and $A \cap B = \varnothing$, then $d(A \cup B) = d(A) + d(B)$.

https://doi.org/10.1017/S0266267118000536 Published online by Cambridge University Press

(5) If $A \in \mathcal{C}_d$, then, for any number n, $A + n \in \mathcal{C}_d$ and $d(A) = d(A + n)$, where $A + n = \{x + n \mid x \in A\}$.
(6) The set of even numbers has density 1/2, or more generally, the set of numbers that are divisible by $m < \infty$ has density $1/m$.

Notice that d is not defined for *all* subsets of $\mathbb{N}$ ($\mathcal{C}_d$ is not a field of natural numbers). We hence seek to extend d to a finitely additive probability measure μ so that μ is defined for all subsets of the natural numbers and that μ agrees with d on $\mathcal{C}_d$. The following is an example of such a measure that is often used in the literature.

Example A.1. Let $\{\lambda_n\}$ be a sequence of functions defined on $\mathbb{N}$ such that

$$\lambda_n(i) = \begin{cases} 1/n & \text{if } 1 \leq i \leq n, \\ 0 & \text{if } i > n. \end{cases} \tag{A.3}$$

It can be shown that the $\lambda_n(i)$s converge to a function λ definable for all subsets of N which extends d.[22] What is important to us is that λ exhibits the following properties:

(1) λ is defined for all subsets of $\mathbb{N}$.
(2) $\lambda(\varnothing) = 0$ and $\lambda(\mathbb{N}) = 1$.
(3) λ is finitely additive.
(4) λ is *not* countably additive.
(5) For any $i < \infty$, $\lambda(\{i\}) = 0$.
(6) For any $A \subseteq \mathbb{N}$, if A is finite then $\lambda(A) = 0$; if A is co-finite (i.e. if $\mathbb{N} - A$ if finite) then $\lambda(A) = 1$.
(7) $\lambda(\{2n \mid n \in \mathbb{N}\}) = 1/2$, i.e. the set of even numbers has measure 1/2.
(8) In general, the set of numbers that are divisible by $m < \infty$ has measure $1/m$, that is, $\lambda(\{1m, 2m, 3m, \ldots\}) = 1/m$. As a result of this property, we have that the assignment of μ can be arbitrarily small: for any $\lambda > 0$, there exists some n such that the set of numbers that are divisible by n has measure $1/n < \varepsilon$. ◁

Appendix B. Adams' money pump

Adams (1962) showed that there are scenarios in which the failure of countable additivity leads to a money pump. More precisely, Adams' example presents a betting situation where a (Bayes) rational gambler is justified in accepting, with a small fee, each bet of a sequences of bets, but the acceptance of all the bets leads to sure loss. For completion, I include a variant of this example here.

Example B.1. Let $S = \mathbb{N}$, $X = [-1,1]$, and let the identity function $u(x) = x$ be the utility function on X. Let λ be the finitely but not countably additive measure on positive integers given in Example A.1 and let η be the countably additive probability measure on S defined by

$$\eta(n) = \tfrac{1}{2^n} \quad \text{for all } n \in S.$$

Define subjective probability μ to be such that

$$\mu(n) = \frac{\lambda(n) + \eta(n)}{2} \quad \text{for all } n \in S. \tag{B.1}$$

The following is a list of simple properties of μ.

(1) μ is a finitely but not countably additive probability measure.

[22]See Rao and Rao (1983: Theorem 3.2.10). Hrbacek and Jech (1999: Ch. 11) contains a set-theoretic construction.

https://doi.org/10.1017/S0266267118000536 Published online by Cambridge University Press

Table B.1. Gamble g_n

	$B_n = \{n\}$	$S - B_n$
g_n	$1/2^{n+1} - r$	$1/2^{n+1}$

(2) For any $n < \infty$,

$$\mu(n) = \frac{0 + \eta(n)}{2} = \frac{1}{2^{n+1}}.$$

(3) $\mu(S) = 1$, whereas

$$\sum_{i=1}^{\infty} \mu(i) = \frac{\sum_{i=1}^{\infty} \lambda(i) + \sum_{i=1}^{\infty} \eta(i)}{2} = \frac{0+1}{2} = \frac{1}{2}.$$

(4) For any finite $E \subseteq S$,

$$\mu(E) = \frac{0 + \eta(E)}{2} = \sum_{i \in E} \frac{1}{2^{i+1}}.$$

Now, for each n, consider gamble g_n with payoff described as in Table B.1. That is to say, g_n pays $1/(2^{n+1})$ no matter which state obtains, but will cost the gambler $r \in \left(\frac{1}{2}, 1\right)$ in the event of $B_n = \{n\}$. Since $r < 1$, it is easy to calculate that, for each n, g_n has a positive expected return:

$$U[g_n] = \frac{1}{2^{n+1}} \cdot \left(\frac{1}{2^{n+1}} - r\right) + \left(1 - \frac{1}{2^{n+1}}\right) \cdot \frac{1}{2^{n+1}} = \frac{1}{2^{n+1}} \cdot (1 - r).$$

Hence, a (Bayes) rational gambler should be willing to pay a small fee ($<U[g_n]$) to accept each gamble. However, the acceptance of all gambles leads to sure loss no matter which number eventually transpires. To see this, note that, for any given number m, gamble g_n pays

$$g_n(m) = \frac{1}{2^{n+1}} - r \cdot \chi_{B_n}(m)$$

But, the joint of all g_ns yields

$$\sum_{i \in S} g_i(m) = \sum_{i \in S} \left[\frac{1}{2^{i+1}} - r \cdot \chi_{B_i}(m)\right] = \frac{1}{2} - r < 0.$$

That is to say, for each possible outcome $m \in S$, the expected value of getting m from accepting all the gambles g_ns is negative. ◁

Appendix C. Savage's postulates

Let S be the set of states of the world and X the set of consequences. $\mathcal{F}$ is a σ-algebra equipped on S. A Savage act f is a function mapping from S to X. An act is said to be *constant* with respect to consequence $a \in X$, denote by $\mathfrak{c}_a$, if $\mathfrak{c}_a(s)$ for all $s \in S$. Define the *combination* of acts f and g with respect to an event E (a set of states), written $f|E + g|\overline{E}$, to be such that:

$$(f|E + g|\overline{E})(s) =\text{Df} \begin{cases} f(s) & \text{if } s \in E \\ g(s) & \text{if } s \in \overline{E}, \end{cases}$$

where $\overline{E} = S - E =$ the complement of E.

https://doi.org/10.1017/S0266267118000536 Published online by Cambridge University Press

Definition (conditional preference). Let E be some event, then, given acts $f, g \in \mathcal{A}$, f is said to be weakly preferred to g *given* E, written $f \succcurlyeq_E g$, if, for *all* pairs of acts $f', g' \in \mathcal{A}$,

(1) f agrees with f' and g agrees with g' on E, and
(2) f' agrees with g' on $\overline{E}$

imply

$$f' \succcurlyeq g'.$$

That is, $f \succcurlyeq_E g$ if, for all $f', g' \in \mathcal{A}$,

$$\left.\begin{array}{ll} f(s) = f'(s),\ g(s) = g'(s) & \text{if } s \in E \\ f'(s) = g'(s) & \text{if } s \in \overline{E}. \end{array}\right\} \Longrightarrow f' \succcurlyeq g'. \tag{C.1}$$

Definition (Null events). An event E is said to be a *null* if, for any acts $f, g \in \mathcal{A}$,

$$f \succcurlyeq_E g. \tag{C.2}$$

That is, an event is null if the agent is indifferent between any two acts *given* E. Intuitively speaking, null events are those events such that, according to the agent's beliefs, the possibility that they occur can be ignored.

Savage's Postulates.

P1: $\succcurlyeq$ is a weak order (complete preorder).
P2: For any $f, g \in \mathcal{A}$ and for any $E \in \mathcal{B}$, $f \succcurlyeq_E g$ or $g \succcurlyeq_E f$.
P3: For any $a, b \in X$ and for any non-null event $E \in \mathcal{B}$, $\mathfrak{c}_a \succcurlyeq_E \mathfrak{c}_b$ if and only if $a \geq b$.
P4: For any $a, b, c, d \in C$ satisfying $a \succcurlyeq b$ and $c \succcurlyeq d$ and for any events $E, F \in \mathcal{B}$, $\mathfrak{c}_a|E + \mathfrak{c}_b|\overline{E} \succcurlyeq \mathfrak{c}_a|F + \mathfrak{c}_b|\overline{F}$ if and only if $\mathfrak{c}_c|E + \mathfrak{c}_d|\overline{E} \succcurlyeq \mathfrak{c}_c|F + \mathfrak{c}_d|\overline{F}$.
P5: For some constant acts $\mathfrak{c}_a, \mathfrak{c}_b \in \mathcal{A}$, $\mathfrak{c}_b \succ \mathfrak{c}_a$.
P6: For any $f, g \in \mathcal{A}$ and for any $a \in C$, if $f \succ g$ then there is a finite partition $\{P_i\}_{i=1}^n$ such that, for all i, $\mathfrak{c}_a|P_i + f|\overline{P_i} \succ g$ and $f \succ \mathfrak{c}_a|P_i + g|\overline{P_i}$.
P7: For any event $E \in \mathcal{B}$, if $f \succcurlyeq_E \mathfrak{c}_{g(s)}$ for all $s \in E$ then $f \succcurlyeq_E g$.

Yang Liu is a Research Fellow in the Faculty of Philosophy at the University of Cambridge. Before he started teaching at Cambridge, he received his doctorate in philosophy from Columbia University. His research interests include logic, foundations of probability, Bayesian decision theory, and philosophy of artificial intelligence. Liu is currently a Leverhulme Early Career Fellow, working on a project titled 'Realistic Decision Theory'. More information is available on his website at http://yliu.net.

Cite this article: Liu Y (2020). Countable additivity, idealization, and conceptual realism. *Economics & Philosophy* **36**, 127–147. https://doi.org/10.1017/S0266267118000536

https://doi.org/10.1017/S0266267118000536 Published online by Cambridge University Press