Aggregating infinitely many probability measures

Frederik Herzberg

Abstract. The problem of how to rationally aggregate probability measures occurs in particular (i) when a group of agents, each holding probabilistic beliefs, needs to rationalise a collective decision on the basis of a single 'aggregate belief system' and (ii) when an individual whose belief system is compatible with several (possibly infinitely many) probability measures wishes to evaluate her options on the basis of a single aggregate prior via classical expected utility theory (a psychologically plausible account of individual decisions).

We investigate this problem by first recalling some negative results from preference and judgment aggregation theory which show that the aggregate of several probability measures should not be conceived as the probability measure induced by the aggregate of the corresponding expected-utility preferences. We describe how McConway's (*Journal of the American Statistical Association*, vol. 76, no. 374, pp. 410–414, 1981) theory of probabilistic opinion pooling can be generalised to cover the case of the aggregation of infinite profiles of finitely-additive probability measures, too; we prove the existence of aggregation functionals satisfying responsiveness axioms *à la* McConway plus additional desiderata even for infinite electorates. On the basis of the theory of propositional-attitude aggregation, we argue that this is the most natural aggregation theory for probability measures.

Our aggregation functionals for the case of infinite electorates are neither oligarchic nor integral-based and satisfy (at least) a weak anonymity condition. The delicate set-theoretic status of integral-based aggregation functionals for infinite electorates is discussed.

Key words: probabilistic opinion pooling, general aggregation theory, Richard Bradley, multiple priors, Arrow's impossibility theorem, Bayesian epistemology, society of mind, finite anonymity, ultrafilter, measure problem, non-standard analysis.

Journal of Economic Literature classification: D71, D81, C11

2010 Mathematics Subject Classification: 91B06, 91B10, 03H05, 03E35, 00A30, 60A05

Munich Center for Mathematical Philosophy, Ludwig Maximilian University of Munich, Geschwister-Scholl-Platz 1, D-80539 Munich, Germany.

Center for Mathematical Economics, Bielefeld University, Universitätsstraße 25, D-33615 Bielefeld. E-mail: fherzberg@uni-bielefeld.de.

1. Introduction

This paper studies the aggregation of infinitely many probability measures. Depending on one's disciplinary outlook, the question may be either of interest in its own right or need further motivation. The rest of the introduction and the following Section 2 is meant to provide such motivation; these initial parts can be ignored by readers who find the problem of sufficient interest in itself.

Formal epistemologists are now increasingly following the decision-theoretic paradigm of recent decades by weakening the strict, classical Bayesian assumption of rational agents being endowed with a single subjective probability measure. (Subjective probability measures are generally referred to as *priors* in the decision-theoretic literature, even when learning by conditioning is very rarely studied, whence there are, strictly speaking, neither prior nor posterior probability measures in the statistical sense; we submit to this terminological convention.) Richard Bradley [**3**], for example, entertains the possibility of rational agents having belief systems that are compatible with several subjective probability measures — as opposed to a single unique one. This appears, at first sight, to be a step in the direction of the recent decision-theoretic literature on multiple priors.

However, this is not the case — due to the different perspectives of epistemologists on the one hand and decision theorists on the other. There are at least two ways in which the aggregation of probability measures *per se* — as opposed to their induced preferences — is of great importance from an epistemological point of view and perhaps also of some interest from a decision-theoretic vantage point

The first comes from social epistemology. Consider a group whose members hold different belief systems, each encoded by a subjective probability measure, and face a collective decision, i.e. they need to choose one of several social alternatives. Suppose this group wants to ensure that their decision is rationally defensible given some belief system which in a reasonable sense aggregates their individual belief system. Let us assume, for simplicity, that the individuals differ only with respect to their beliefs while sharing a common utility function over final outcomes. A good solution to their problem would be to look for a way of rationally aggregating the probability measures describing their individual belief systems and afterwards choosing a social alternative which maximises expected utility with respect to the aggregate probability measure. (Cases similar to this, however with additional serious complications, feature in a forthcoming paper by Bradley, Dietrich and List [**4**].)

Another area in which the aggregation of probability measures *per se* becomes important is a new, epistemologically and psychologically informed account of individual decisions. This account is motivated by theories from contemporary psychology according to which the human mind is to be understood as a composition of more elementary mental agents, viz. in terms of a "society of mind" (Minsky [**44**]), a "multimind" (Ornstein [**45**]), or an "internal family system" (Schwartz [**50**]); it has been formalised and introduced into the epistemological discussion by Richard Bradley [**3**].

Technically speaking, this new account of human decision making attempts to give a precise description of the decision process of an individual whose belief system is compatible with several probability measures: Bradley [**3**] suggests that every decision of an agent with multiple subjective probability measures is preceded by an "aggregation" of those priors — which takes her, temporarily, to a new, aggregate prior, i.e. a new probabilistically consistent belief system; in order to make a decision, she will then evaluate the options available to her using the classical expected utility criterion with respect to the temporary, aggregate prior.[1] After the decision has been made, she returns to her previous epistemic state encoded by a whole set of subjective probability measures. Note that this set of subjective probability measures will in general not be finite — e.g. when it is the set of all probability measures that are consistent with certain conditional probability assignments or that satisfy certain upper or lower bounds when evaluated at certain events ("interval probabilities"). Therefore, epistemologists who wish to follow Richard Bradley [**3**] in his proposal to integrate insights from contemporary psychology into epistemology should seriously consider the problem of aggregating *infinite* profiles of probability measures.

Of course, the classical decision theory for multiple subjective probability measures (*multiple priors*), as in Gilboa and Schmeidler [**18**], is not — at least not without substantial, quite probably philosophically questionable, detours — applicable in these situations: Whilst one could associate each probability measure in the set of priors with an individual and thereby view a maxmin expected-utility preference ordering as an aggregate preference ordering, the corresponding "aggregator" would aggregate preferences derived from priors, not priors themselves. Therefore, this theory is — notwithstanding its mathematical elegance and also in spite of its manifold practical use, e.g. in mathematical finance (cf. e.g. Riedel [**46**]) — not satisfactory for the epistemologically motivated purposes of the present paper.

The paper is structured as follows. Section 2 motivates the problem of aggregating subjective probability measures from a revised Bayesian perspective. This section can be skipped by readers who are only interested in the technical aggregation problem itself, which will be studied in Section 3. We will see that there is indeed a relatively natural and decision-theoretically defensible way of aggregating probability measures — via a generalised aggregation theory for infinite profiles of probability measures along the lines of McConway [**43**] (and Arrow [**2**]). Moreover, this existence question is not trivial, as certain natural candidates like oligarchic aggregation or aggregation via integrals on the electorate are not feasible. The formal details and proofs are given in Sections 4 and 5, respectively. Section 6 concludes.

2. Motivation: Bayesianism with multiple priors

2.1. Single- vs. multiple-prior Bayesianism and decision theory.
In classical Bayesian contributions to epistemology and decision theory it is often assumed that the degrees of belief of a rational individual can be

---

[1]An agent will always have several probability measures if her belief system only entails a partial probability assignment, as for instance in Jeffrey's "radical probabilism" [**29**].

given by a single probability measure: A rational individual, it is tacitly or explicitly assumed in this literature, makes judgments about all propositions in some algebra and does so by assigning a precise real number, its degree of belief, to each of them. Let us call this thesis — for reasons which will shortly become clear — *single-prior Bayesianism*. For philosophical simplicity, let us also assume that all credences are reported explicitly by the individual, rather than implicitly (through betting or the like).

Of course, this is a very strong assumption. To be sure, formal epistemologists have offered numerous arguments why the degrees of belief of a rational agent should satisfy the axioms of probability theory (at least with finite additivity), if they are precise (e.g. Joyce [**30**], Easwaran and Fitelson [**14**], Fitelson and McCarthy [**17**], Leitgeb and Pettigrew [**37, 38**], Wedgwood [**54**]). These arguments, however, do not establish that the system of propositions to which a rational agent assigns degrees of belief must be an algebra, nor that a rational agent must always assign *precise* degrees of belief.

Indeed, rational agents may subscribe to a set $\mathcal{S}$ of precise assignments of *conditional* degrees of belief (i.e. assignments of real values to conditional events $\langle A|B\rangle$, where A, B belong to some algebra $\mathcal{A}$ of propositions of which the agent is aware) which is in general too small to derive precise degrees of belief even for all propositions which occur in the conditional degree-of-belief assignments in $\mathcal{S}$. For instance, a rational agent may (for symmetry reasons, say) assign a conditional degree of belief of $1/2$ to proposition A *given* B, but at the same time may not be able to assign a precise degree of belief to either A or B. Instead, she would deem a whole, in general: infinite, set of probability measures compatible with her beliefs. Let us denote the thesis that the belief system of a rational agent corresponds to a set of probability measures by *multiple-prior Bayesianism*.

Additional reasons for relaxing the assumption of single-prior Bayesianism, i.e. that the beliefs of a rational agent can always be encoded by a single probability measure, are provided by decision theory. More than ninety years ago, F. Knight [**33**] already distinguished two kinds of uncertainties to which economic agents may be subject: Sometimes, economic agents are merely uncertain about the exact value of a certain economic variable while being quite certain about the probability distribution of that variable. This kind of uncertainty is called *risk* or *first-order uncertainty*. But there are also situations in which the economic agents do not even know the exact probability distribution of some non-deterministic economic variable, but consider several, perhaps rather different, probability distributions possible. This kind of uncertainty is referred to as *second-order risk* or *second-order uncertainty* or *Knightian uncertainty* and is empirically observable (cf. e.g. the famous Ellsberg paradox [**15**]). Adopting the statistical terminology of prior and posterior probability distributions, the phenomenon of second-order uncertainty is sometimes also expressed by the term *multiple priors* or *ambiguity* (of the prior).

For single-prior Bayesianism, there is an obvious decision criterion: An act is preferable to another act if and only if the expected utility of the

outcome of the first act under the single prior is higher than the expected utility of the outcome of the second act. This is not only an intuitively appealing, but also a particularly *rational* decision criterion in a rigorous theoretical sense: The expected-utility theorems of de Finetti [**8, 9, 10, 11**], von Neumann and Morgenstern [**53**] and, as a synthesis, Savage [**49**] show that any preference relation among acts (i.e. maps from a certain set, whose elements are called *states of the world*, to another set, whose elements are called *outcomes*) which satisfies certain rationality postulates can be derived from an expected-utility criterion for some probability distribution on the set of states of the world and some utility function defined on the set of outcomes.

2.2. Options for a Bayesian decision theory with multiple priors. Among the strengths of single-prior Bayesianism — the thesis that the belief system of a rational agent can always be adequately captured by a single unique probability measure — is its association with a both theoretically and practically very appealing decision criterion. But what would be appropriate decision criteria under the hypothesis of multiple-prior Bayesianism — i.e. the thesis that rational agents may have belief systems compatible with several subjective probabiltity measures ("Knightian [**33**] uncertainty")? This depends, unsurprisingly, crucially on what we mean by "appropriate" in this context. Two approaches have to be distinguished.

We might either conceive of multiple-prior Bayesian agents as rational decision makers, whence "appropriate" reduces to a notion of decision-theoretic rationality. If we were to choose that approach, we would look for decision criteria that have some intuitive appeal and at the same time can be rationalised in the same manner as the expected-utility criterion has been rationalised by de Finetti [**8, 9, 10, 11**], von Neumann and Morgenstern [**53**] and ultimately Savage [**49**]. (An additional criterion might perhaps be the practical implementability of a decision criterion.)

Among the most famous results in this direction are the rationalisation of the maxmin expected-utility decision criterion by Gilboa and Schmeidler [**18**] and its generalisation to the decision criterion of maxmin expected-utility with additive penalty (encoding the relative unlikelihood of certain priors) by Maccheroni, Marinacci and Rustichini [**42**].

However, we might also look at our multiple-prior Bayesian agents (agents who are subject to Knightian uncertainty) from a more epistemologically motivated perspective. This is what Richard Bradley [**3**] suggested in a recent colloquium talk. In line with some of the recent psychological literature (for instance, Minsky [**44**], Ornstein [**45**], or Schwartz [**50**]), Bradley conceives of multiple-prior Bayesian agents as complex — in extreme cases schizophrenic — personalities, composed of simpler sub-personalities (called "opinionated avatars" by Bradley) corresponding to each prior that is compatible with the agent's belief system. When the time is ripe for a decision, the complex personality *aggregates* (Bradley) her different probability measures into a single, temporary subjective probability measure and then choose the optimal act based on expected utility; after the decision has been made, she continues to be the complex personality that she was before.

In other words, on Bradley's [3] account, the preferences of a multiple-prior Bayesian agent (on which the decision is based) are not deduced directly from her set of priors — as would be the case if the agent were to use the maxmin expected-utility decision criterion. Instead, her preferences are derived, at least for this specific decision, via classical expected-utility theory, from a single, temporary subjective probability measure which has been constructed using some aggregation rule from her set of priors — even though the aggregation rule might vary, thus allowing for time-inconsistency. Hence, at least from an epistemological perspective, the question now remains how one can merge ('aggregate') probability measures; behaviourally, only the aggregate prior is revealed.

3. The aggregation of probability measures

In principle, there are multiple possibilities for the 'aggregation' of several priors into a single one. One's choice will have to depend on how closely one wants to follow some established notion of orthodoxy for aggregation.

If one were to take the requirement of an orthodox aggregation theory very literally, one might recall that Arrow's [2] aggregation theory is *the* classical microeconomic theory of aggregation. Since its domain is the aggregation of preferences, one would have to convert the priors first of all into preference orderings; the canonical way to achieve this is via expected-utility theory. This will then prompt the question under which circumstances Arrovian aggregation of expected-utility preferences is possible.

One might even go one step further (than Bradley [3]) in the direction of a *society of mind* (Minsky [44]) psychology of decisions: What if the decision maker is composed of agents that are themselves conglomerates of even more elementary agents who face some second-order uncertainty, yet have very simple utility functions. Then, under some rationality constraints, one may assume that each of the agents that constitute the decision maker has maxmin expected-utility preferences or at least preferences from one of the larger classes of ambiguous preferences, such as the variational preferences of Maccheroni, Marinacci and Rustichini [42] or even the MBA preferences introduced by Cerreia-Vioglio, Ghirardato, Maccheroni, Marinacci and Siniscalchi [6]. The above question then becomes whether Arrovian aggregation of such generalisations of expected-utility preferences is possible.

As it turns out, this is in general not the case: Whenever one wants to aggregate a profile of variational preferences from a sufficiently rich class (e.g. the set of all expected-utility preferences for some set of states of the world or the set of all multiple-prior preferences) into a single variational preference ordering (e.g. an expected-utility preference ordering on that set of states of the world), there will be no aggregation rule satisfying the analogues of Arrow's responsiveness axioms, as was shown in Herzberg [23, 22]. Note that these impossibility statements can be established both for the case of profiles of a given finite length (the analogue of Arrow's [2] theorem) and for the case of profiles of any given infinite length (the analogue of Campbell's [5] theorem), using a model-theoretic approach to aggregation theory inspired by Lauwers and Van Liedekerke [35] and systematically

elaborated by Herzberg and Eckert [**25, 26**]. For the special case of the Arrovian finite expected-utility profiles of a given length, this impossibility theorem was proved by Le Breton [**36**]; an impossibility theorem for the non-dictatorial aggregation of expected-utility preferences in a slightly different, yet still very natural setting was established by Hylland and Zeckhauser [**27**].

Thus, one will have to relax somehow the requirement of following a well-established orthodoxy regarding aggregation. Another approach becomes available as soon as one recalls that the problem of merging several priors into a single prior has actually already been studied in the statistical literature: under the heading of probabilistic opinion pooling, in particular McConway's [**43**] characterisation of aggregation (pooling) rules satisfying certain responsiveness axioms as linear averaging rules. (McConway's [**43**] findings have also entered the literature on aggregating expert judgments, e.g. Cooke [**7**]; the aggregation of probability distributions, yet without the imposition of Arrovian responsiveness axioms, has been studied by Lindley [**39**].)

Even though McConway has shown a possibility result and Arrow an impossibility result, the responsiveness axioms for pooling rules imposed by McConway [**43**] are remarkably similar to those of the Arrovian [**2**] social choice literature. In fact, based on these similarities, Dietrich and List [**13, 12**] have formulated a unified framework for preference aggregation, probabilistic opinion pooling and judgment aggregation. Herzberg [**24**] has proposed a unified mathematical methodology for approaching this novel general aggregation theory of Dietrich and List [**13, 12**]; in this setting, one can derive both McConway's [**43**] theorem and a judgment-aggregation analogue of Arrow's theorem [**2**] from a single characterisation theorem has been proposed in Herzberg [**24**]. The basic idea is that in a binary setting, there are fewer aggregation rules anyhow, whence the responsiveness axioms can only be satisfied by aggregation ruler that are projections, while in the probabilistic setting, the space of possible aggregation rules is much larger and analogous responsiveness axioms can be satisfied by linear aggregation rules.

Understandably, the probabilistic opinion pooling literature is only concerned with finite sets of (σ-additive) priors. Since the set of priors of multiple-prior Bayesian agents (agents facing Knightian uncertainty) will in general be infinite sets of finitely-additive priors, we are looking for an appropriate generalisation of McConway's theorem. Our desideratum is a theorem which proves the existence of aggregation rules for infinite profiles of priors that satisfy Arrow's responsiveness axioms.

In fact, it is not too difficult to give such an existence proof, at least if one does not insist on the σ-additivity of the aggregate prior (which also permits to study the aggregation of profiles of finitely-additive probability measures).

We adopt essentially the same setup as in McConway [**43**], with three modifications. First, no σ-additivity is required of the probability measures. Secondly (WSFP, SSFP, or ZPP) consensus functions are now called (independent, systematic, or unanimity-preserving — respectively)

aggregation functionals, in order to connect the result to general aggregation theory. Thirdly and most importantly, we allow for an infinite electorate N.

In the remainder of the paper, we shall discuss the framework just described (including the assumptions on the aggregation operators) and prove the following results, assuming that the underlying set of states of nature has at least three elements:

(1) An aggregation functional is systematic if and only if it is both independent and unanimity-preserving.
(2) If the electorate is finite, then an aggregation functional is systematic if and only if the aggregation functional can be reduced to weighted averaging of probabilities.
(3) If the electorate is infinite, then there again exist aggregation functionals that are systematic but cannot be derived from the probability assignments of any finite subset of the electorate.
(4) The existence of such aggregation functionals is non-trivial because an obvious class of candidates for such aggregation functionals, viz. integrals with respect to a σ-additive measure on the (power-set of the) electorate, generically is empty if N is uncountable.[2]

Thus, non-trivial systematic aggregation of priors is possible even for infinite electorates. Parts 1 and 2, of course, are merely a straightforward adaptation of McConway's famous results [**43**, Theorems 3.2 and 3.3] to a setting in which σ-additivity of priors is not assumed and which also admits with profiles of infinite length.

For part 3 of the Theorem, one has to find systematic aggregation functionals for infinite profiles of priors, which have hitherto not been constructed. Our construction uses Robinson's [**47, 48**] nonstandard analysis through an ultrapower of the reals with respect to some non-principal ultrafilter on the cardinality of the set of priors (i.e. the profile length). Very roughly speaking, the ultrafilter can be seen as a device of picking an accumulation point of a bounded sequence (such as the profile of all probabilities assigned to a particular event by individuals in the electorate), and the use of nonstandard analysis permits (a) calculating with this accumulation point as if it were an ordinary sequence element (i.e. just an ordinary real between 0 and 1), and (b) extracting accumulation points in a uniform manner for all sequences (i.e. all profiles of probabilities). In particular, this construction satisfies only a weak anonymity concept, viz. finite anonymity, but not bounded or even strong anonymity, as was shown by Fey [**16**].

Part 4 merely assembles various results from axiomatic set theory on the so-called *measure problem*. The problem with integral-based aggregation functionals is that the function $i \mapsto P_i(A)$ which assigns to each individual i the subjective probability of i for a given event A, given some profile $\underline{P}$, need not be measurable with respect to a fixed σ-algebra on an uncountable electorate N. In general, it can only be assumed to be measurable with respect to the power-set σ-algebra. However, the existence of σ-additive probability measures on the power-set of an uncountable set N is an intricate

---

[2]I am very grateful to the Associate Editor for the suggestion to look at this question.

set-theoretic problem; the existence of such a measure cannot be proved in classical mathematics and for some cardinalities of N even refuted.

4. A generalised aggregation theory for probability measures *à la* McConway and Arrow

The possibility of aggregating finitely many probability measures while satisfying responsiveness axioms similar to Arrow's [2] has been known for more than three decades: Any rule which takes the weighted average of the individual probability measures already has those desirable properties, and there are no other rules satisfying those desiderata (McConway [43]).

What has not been treated in the existing literature — at least to the present author's knowledge — is the aggregation of infinitely many probability measures. We will now show that one can aggregate an infinite number of subjective finitely additive probability measures in a manner that satisfies McConway's [43] Arrovian responsiveness axioms. It is enough to form an ultrapower of the sequence of the subjective finitely additive probability measures, with respect to an ultrafilter over the infinite set of individuals, and then push it down to a standard finitely additive probability measure by composing it event-wise with the standard part operator (as in the Loeb [40] measure construction, but without requiring σ-additivity and thus without invoking a saturation principle).

4.1. Formal framework. In the following, we present a framework for probabilistic opinion pooling which is very similar to that of the classical paper by McConway [43]. Our framework is more general in that it allows for infinite electorates as well (and is concerned with the aggregation of finitely additive probability measures). Our terminology reflects the formal similarities between judgment aggregation and probabilistic opinion pooling which have already been observed by Dietrich and List [13] (for a more formal treatment, see Herzberg [24]).

Let Ω be a set of possible worlds and let Σ be the set of all algebras on Ω.[3] For all $\mathcal{A} \in \Sigma$, let $\Delta(\mathcal{A})$ be the set of finitely additive probability measures defined on $\mathcal{A}$.

Let N be a finite or infinite set, called the *electorate*. An *aggregation functional* is a map F with domain Σ such that $F(\mathcal{A}) : \Delta(\mathcal{A})^N \to \Delta(\mathcal{A})$ for all $\mathcal{A} \in \Sigma$. (McConway [43] calls our aggregation functionals "classes of consensus functions", denoted by $\mathcal{C}$.) Elements $\Delta(\mathcal{A})^N$ (for any $\mathcal{A} \in \Sigma$) are called *profiles* and are typically denoted $\underline{P}$. Given some profile $\underline{P} = (P_i)_{i \in N} \in \Delta(\mathcal{A})^N$ (for any $\mathcal{A} \in \Sigma$), we denote by $\underline{P}(A)$, for any $A \in \mathcal{A}$, the sequence $(P_i(A))_{i \in N} \in [0,1]^N$. Furthermore, we denote by $\underline{0}$ and $\underline{1}$ the N-sequences consisting only of 0's and 1's, respectively. Note that whenever $\mathcal{A} \neq \mathcal{B}$, $\Delta(\mathcal{A}) \cap \Delta(\mathcal{B}) = \varnothing$ and therefore the domains of $F(\mathcal{A})$ and $F(\mathcal{B})$ are disjoint. We shall therefore usually drop the first argument, writing $F(\underline{P})$ instead of $F(\mathcal{A})(\underline{P})$ for all $\mathcal{A} \in \Sigma$.

---

[3]One could simplify the framework, in departure from McConway's [43] original paper, by just considering a single algebra, thereby ensuring that all aggregation functionals have the same domain. Since all algebras are, up to isomorphism, subalgebras of power-set algebras (due to Stone's [51] representation theorem), this framework would be less general than ours.

4.2. Properties of aggregation functionals. A very natural condition on any aggregation functional ("class of consensus functions") is the *unanimity preservation principle* (in McConway's [**43**] terminology: "Zero Probability Property"), which demands that the aggregate probability of some event should be zero if all voters assign zero to it:[4]

DEFINITION 1. *An aggregation functional F is* unanimity-preserving/ZPP *if and only if for all $\mathcal{A} \in \Sigma$, $\underline{P} \in \Delta(\mathcal{A})^N$ and $A \in \mathcal{A}$, one has $F(\underline{P})(A) = 0$ whenever $P_i(A) = 0$ for all $i \in N$.*

Another natural, but perhaps somewhat less compelling condition is to demand that the aggregate probability of some event should depend on nothing else than that event and the sequence of probabilities assigned to that event by the voters (in McConway's [**43**] terminology: "Weak Setwise Function Property"):

DEFINITION 2. *An aggregation functional F is* independent/WSFP *if and only if there exists a function $G : \left(2^\Omega \setminus \{\varnothing, \Omega\}\right) \times [0,1]^N \cup \{(\varnothing, \underline{0})\} \cup \{(\Omega, \underline{1})\} \to [0,1]$ such that for all $\mathcal{A} \in \Sigma$, $\underline{P} \in \Delta(\mathcal{A})^N$ and $A \in \mathcal{A}$,*

$$(1) \qquad\qquad F(\underline{P})(A) = G(A, \underline{P}(A)).$$

The independence property can be traced back to the Arrovian literature where it appears as the requirement of "independence of irrelevant alternatives". In the judgment aggregation literature, it is commonly known as independence *tout court.* The idea is that the social judgment with respect to some proposition (encoding, e.g., a preference of one alternative over another) should only depend on the individuals' attitudes towards that particular proposition. Of course, such a requirement is only plausible if the agenda admits some kind of separability, so that the propositions about which social judgments are formed enjoy some degree of mutual independence themselves.

It is clear that $G(\Omega, \underline{1}) = 1$ and $G(\varnothing, \underline{0}) = 0$ whenever G satisfies Equation (1) for some aggregation functional F (because $F(\underline{P})$ is, by assumption on F, always a finitely-additive probability measure). Moreover, the extension of the notion of independence/WSFP would not change if one replaced the domain of G by $2^\Omega \times [0,1]^N$. This, however, would introduce additional notational difficulties in the proof.

Finally, an even stronger notion than independence would be to require that the aggregate probability of some event should depend on nothing else but the sequence of probabilities assigned to that event by the voters (in McConway's [**43**] terminology: "Strong Setwise Function Property"):

DEFINITION 3. *An aggregation functional is* systematic/SSFP *if and only if there exists a function $f : [0,1]^N \to [0,1]$ such that for all $\mathcal{A} \in \Sigma$,*

---

[4]An alternative name for this concept might be the Pareto principle. However, the Pareto principle is commonly associated with aggregation of preferences as opposed to beliefs. (Even though one might argue that aggregation of probabilistic beliefs cannot be separated from aggregation of utilities, cf. Hylland and Zeckhauser [**27**].) The desideratum of unanimity preservation is applicable to epistemic as well as economic aggregation problems.

$\underline{P} \in \Delta(\mathcal{A})^N$ *and* $A \in \mathcal{A}$,

$$(2) \qquad\qquad F\left(\underline{P}\right)(A) = f\left(\underline{P}(A)\right).$$

The notion of systematicity is known in the preference aggregation literature as "neutrality". It can be seen as stronger version of independence, demanding that aggregation procedures be blind to the content of the proposition about which individual judgments are aggregated.

From the judgment aggregation literature, it is well-known that for sufficiently complex agendas, systematicity and indepencence are actually equivalent (for a consise proof, cf. e.g. Eckert and Klamler [**32**]). A similar finding holds in the setting of McConway's [**43**] probabilistic opinion pooling. This result of McConway's [**43**, Theorem 3.2] can easily be generalised to infinite electorates, as follows. (Because the original proof is only for finite electorates and contains several misprints, we shall provide a full proof, even though it is very close to McConway's.)

THEOREM 4 ("Unanimity-preservation & Independence = Systematicity"). *Suppose* card$(\Omega) \geq 3$. *Then, an aggregation functional is both unanimity-preserving/ZPP and independent/WSFP if and only if it is systematic/SSFP.*

4.3. Existence and characterisation of aggregation functionals. For the formal statement of our main theorem, we need the notion of an oligarchy, which is a finite subset of an infinite electorate which uniquely determines the aggregation functional by means of weighted averaging.

DEFINITION 5. *An aggregation functional is an* oligarchy *if and only if there exists a finite proper subset* $M \subset N$ *and some function* $h : [0,1]^M \to [0,1]$ *such that* $F(\underline{P})(A) = h\left((P_i(A))_{i \in M}\right)$ *for all* $\mathcal{A} \in \Sigma$, $\underline{P} \in \Delta(\mathcal{A})^N$ *and* $A \in \mathcal{A}$.

Beyond non-oligarchy, the stronger notion of finite anonymity is a desirable property of aggregation functionals for infinite electorates:[5]

DEFINITION 6. *An aggregation functional is* finitely anonymous *if and only if for each permutation* $\pi : N \to N$ *that is constant on a co-finite subset of* N, *one has* $F(\underline{P})(A) = F\left(\left(P_{\pi(i)}\right)_{i \in N}\right)(A)$ *for all* $\mathcal{A} \in \Sigma$, $\underline{P} \in \Delta(\mathcal{A})^N$ *and* $A \in \mathcal{A}$.

If there exists some probability measure μ on the whole power-set of N, then $i \mapsto P_i(A)$ will be a bounded measurable function for all $\mathcal{A} \in \Sigma$, $\underline{P} \in \Delta(\mathcal{A})^N$ and $A \in \mathcal{A}$. Hence one may define in such a setting a systematic/SSFP aggregation functional by $F(\underline{P})(A) = \int_N P_i(A)\mu(\mathrm{d}i)$ for all $\mathcal{A} \in \Sigma$, $\underline{P} \in \Delta(\mathcal{A})^N$ and $A \in \mathcal{A}$. The existence of such a μ, however, becomes a profound set-theoretic problem if N is uncountable. The consequence of this will be seen in the subsequent characterisation theorem for aggregation functionals. Let us, for simplicity, call aggregation functionals *continuous linear* if and only if they admit an integral representation of the above form.

---

[5]I thank the Associate Editor for suggesting this.

DEFINITION 7. *An aggregation functional is* continuous linear *if and only if there exists a σ-additive measure $\mu : 2^N \to [0,1]$ such that for all $\mathcal{A} \in \Sigma$, $\underline{P} \in \Delta(\mathcal{A})^N$ and $A \in \mathcal{A}$,*

$$F(\underline{P})(A) = \int_N P_j(A)\mu(\mathrm{d}j).$$

REMARK 8. *Any continuous linear aggregation functional is systematic/SSFP.*

LEMMA 9. *If N is countably infinite, then no continuous linear aggregation functional is finitely anonymous.*

THEOREM 10 (Existence of systematic aggregation functionals). *Suppose $\mathrm{card}(\Omega) \geq 3$.*

- *If N is finite, then an aggregation functional F is systematic/SSFP if and only if it is continuous linear, and there exist non-oligarchic aggregation systematic/SSFP functionals.*
- *If N is countably infinite, there exist non-oligarchic aggregation functionals that are continuous linear, but also non-oligarchic systematic/SSFP aggregation functionals that are finitely anonymous.*
- *If N is uncountably infinite, then there exist aggregation functionals that are systematic/SSFP, finitely anonymous and non-oligarchic.*
- *If N is uncountably infinite, the existence of finitely anonymous[0] continuous linear aggregation functionals cannot be proved from Zermelo–Fraenkel set theory with the Axiom of Choice (**ZFC**).*
- *If the cardinality of N is the least uncountable cardinal (or any other successor cardinal), there cannot be a continuous linear aggregation functional.*

The proof follows from the following two lemmas in combination with known results from axiomatic set theory. The first lemma is just a slight variation of McConway's [**43**, Theorem 3.3].

LEMMA 11. *If N is finite, then an aggregation functional F is systematic/SSFP if and only if there exists some $\underline{\alpha} \in [0,1]^N$ such that $F(\underline{P})(A) = \sum_{i \in N} \alpha_i P_i(A)$ for all $\mathcal{A} \in \Sigma$, $\underline{P} \in \Delta(\mathcal{A})^N$ and $A \in \mathcal{A}$.*

LEMMA 12. *If N is infinite, then there exist aggregation functionals that are systematic/SSFP and finitely anonymous, but neither oligarchic nor continuous linear.*

5. Proofs

PROOF OF THEOREM 4. First consider a systematic/SSFP aggregation functional F, and let $f : [0,1]^N \to [0,1]$ be such that for all $\mathcal{A} \in \Sigma$ and every $A \in \mathcal{A}$, Equation (2) holds. Trivially, it is then independent/WSFP. In order to see that it is also unanimity-preserving/ZPP, it is enough to prove that $f(\underline{0}) = 0$. However, again by Equation (2), for any $\mathcal{A} \in \Sigma$ and $\underline{P} \in \Delta(\mathcal{A})^N$,

$$f(\underline{0}) = f(\underline{P}(\varnothing)) = F(\underline{P})(\varnothing) = 0,$$

---

[0]Corrected after publication.

wherein the first equality holds because each P_i is a finitely-additive probability measure and hence $P_i(\varnothing) = 0$ for all $i \in N$, and the last equality holds because $F(\underline{P})$ is a finitely additive probability measure by definition of F.

Now consider an aggregation functional F that is both unanimity-preserving/ZPP and independent/WSFP. Then there exists a function $G : \left(2^\Omega \setminus \{\varnothing, \Omega\}\right) \times [0,1]^N \cup \{(\varnothing, \underline{0})\} \cup \{(\Omega, \underline{1})\} \times [0,1]^N \to [0,1]$ such that Equation (1) holds for all $\mathcal{A} \in \Sigma$ and every $A \in \mathcal{A}$. Because F is unanimity-preserving/ZPP, it is easy to see that we must have

$$(3) \qquad\qquad G(A, \underline{0}) = 0$$

for all $A \subsetneq \Omega$. Let $\underline{P}$ with $\underline{P}(A) = \underline{0}$, so that $\underline{P}(\complement A) = \underline{1}$. Then, $G(\complement A, \underline{1}) = F(\underline{P})(\complement A) = 1 - G(A, \underline{0}) = 1$. Therefore,

$$(4) \qquad\qquad G(B, \underline{1}) = 1$$

for all non-empty $B \subseteq \Omega$.

Consider now two sequences $(a_i)_{i \in N}, (b_i)_{i \in N} \in [0,1]^N$ such that $a_i + b_i \leq 1$ for all $i \in N$. Since $\operatorname{card}(\Omega) \geq 3$, there will be two disjoint non-empty sets $A, B \subsetneq \Omega$ such that $A \cup B \neq \Omega$. Then, there will be a algebra $\mathcal{A}$ such that $A, B \in \mathcal{A}$ and a sequence $\underline{P} \in \Delta(\mathcal{A})^N$ such that for all $i \in N$,

$$P_i(A) = a_i, \qquad P_i(B) = b_i, \qquad P_i(A \cup B) = a_i + b_i.$$

Since $F(\underline{P})$ is a finitely additive probability measure on $\mathcal{A}$ by definition of F, we will have $F(\underline{P})(A \cup B) = F(\underline{P})(A) + F(\underline{P})(B)$, so

$$G(A \cup B, (a_i + b_i)_{i \in N}) = G(A, (a_i)_{i \in N}) + G(B, (b_i)_{i \in N}).$$

In the special case, $(b_i)_{i \in N} = \underline{0}$, we obtain, in light of Equation (3),

$$G(A \cup B, (a_i)_{i \in N}) = G(A, (a_i)_{i \in N}).$$

Therefore (putting $C = A \cup B$), whenever $A \subseteq C \subsetneq \Omega$,

$$(5) \qquad\qquad G(C, \cdot) = G(A, \cdot).$$

(This equation, and structural similar identities in the remainder of this paper, should be understood as an equality on the intersection of the domains of the left-hand side and the right-hand side.)

It is now easy to prove $G(A_1, \cdot) = G(A_2, \cdot)$ for all $A_1, A_2 \subsetneq \Omega$. If A_1 and A_2 have a non-empty intersection, then $G(A_2, \cdot) = G(A_1 \cap A_2, \cdot) = G(A_1, \cdot)$ after applying Equation (5) twice (each time with $A = A_1 \cap A_2$, once with $C = A_2$ and once with $C = A_1$). If A_1 and A_2 are disjoint with $A_1 \cup A_2 \subsetneq \Omega$, then $G(A_2, \cdot) = G(A_1 \cup A_2, \cdot) = G(A_1, \cdot)$ after applying Equation (5) twice (each time with $C = A_1 \cup A_2$, once with $A = A_2$ and once with $A = A_1$). If A_1 and A_2 are disjoint with $A_1 \cup A_2 = \Omega$, then there must be a proper non-empty subset D of either A_1 or A_2. Without loss of generality, assume $\varnothing \neq D \subsetneq A_1$ Then, on the one hand, $G(A_1, \cdot) = G(D, \cdot)$ by Equation (5), and on the other hand $G(A_2, \cdot) = G(A_2 \cup D, \cdot) = G(D, \cdot)$ after noting that $A_2 \cup D \subsetneq A_1 \cup A_2 = \Omega$ and applying Equation (5) twice.

Therefore, if we fix any non-empty $A_1 \subsetneq \Omega$, we will get $G(A, \cdot) = G(A_1, \cdot)$ for all $A \subseteq \Omega$ and we may define $f = f(A_1, \cdot)$. $\qquad\square$

Proof of Lemma 9. If F is continuous linear, say represented by some probability measure μ on N, and N is countably infinite then there must be two individuals i, j such that $\mu\{i\} \neq \mu\{j\}$. Any permutation π that is constant on a co-finite subset of N and satisfies $\pi(i) = j$ and $\pi(j) = i$ is a counterexample to finite anonymity.   □

Proof of Lemma 11. One can literally copy the proof of McConway's main result [**43**, Theorem 3.3], because that proof does not require the σ-additivity of the probability measures in the profile.   □

Proof of Lemma 12. Our proof employs Robinsonian [**47**, **48**] nonstandard analysis.[6] Fix a non-principal ultrafilter $\mathcal{U}$ on N.[7] Then, the ultrapower $\mathbf{R}^N/\mathcal{U}$ will be a non-standard model of the real numbers. Let $^\circ$ denote the standard-part operator on this model of the hyperreals.

Let $\mathcal{A} \in \Sigma$ and $\underline{P} \in \Delta(\mathcal{A})^N$, and define a real-valued set function $F(\underline{P})$ by

$$F(\underline{P}) : \mathcal{A} \to [0, 1], \quad A \mapsto {}^\circ [\underline{P}(A)]_{\sim_\mathcal{U}}.$$

(Since each element of the sequence $\underline{P}(A)$ is between 0 and 1, the ultrapower element $[\underline{P}(A)]_{\sim_\mathcal{U}}$ is a hyperreal between 0 and 1 by Łoś's theorem [**41**], whence its standard part $^\circ [\underline{P}(A)]_{\sim_\mathcal{U}}$ is well-defined and a standard real number between 0 and 1.)

Now, each P_i is finitely additive, addition on the ultrapower $\mathbf{R}^N/\mathcal{U}$ is defined representative-wise and $^\circ$ commutes with addition of limited hyperreals. Therefore, for all disjoint $A, B \in \mathcal{A}$,

$$\begin{aligned} {}^\circ [\underline{P}(A \cup B)]_{\sim_\mathcal{U}} &= {}^\circ [\underline{P}(A) + \underline{P}(B)]_{\sim_\mathcal{U}} = {}^\circ \left([\underline{P}(A)]_{\sim_\mathcal{U}} + [\underline{P}(B)]_{\sim_\mathcal{U}} \right) \\ &= {}^\circ [\underline{P}(A)]_{\sim_\mathcal{U}} + {}^\circ [\underline{P}(B)]_{\sim_\mathcal{U}}, \end{aligned}$$

whence the set function $F(\underline{P})$ defined above is finitely additive. Similarly,

$$ {}^\circ [\underline{P}(\Omega)]_{\sim_\mathcal{U}} = {}^\circ [\underline{1}]_{\sim_\mathcal{U}} = {}^\circ 1 = 1,$$

whence $F(\underline{P})(\Omega) = 1$. Hence, $F(\underline{P})$ is a finitely additive probability measure on $\mathcal{A}$.

We also need to show that F is not an oligarchy. Suppose otherwise, so that there is some $M \subseteq N$ and some function $h : [0, 1]^M \to [0, 1]$ such that $F(\underline{P})(A) = h\left((P_i(A))_{i \in M}\right)$ for all $\mathcal{A} \in \Sigma$, $\underline{P} \in \Delta(\mathcal{A})^N$ and $A \in \mathcal{A}$. On the one hand, one can show that h must be a linear weighted averaging operation, simply by applying the first part of the theorem to the aggregation functional $F^{(M)}$ defined through

$$F^{(M)}(\mathcal{A}) : \Delta(\mathcal{A})^M \to \Delta(\mathcal{A}), \qquad \underline{P} \mapsto h\left((P_i(\cdot))_{i \in M}\right).$$

On the other hand, by our above choice of F, we have

(6) $$ {}^\circ [\underline{\alpha}]_{\sim_\mathcal{U}} = h(\underline{\alpha})$$

for all $\underline{\alpha} \in [0, 1]^N$, in particular for all sequences that converge to zero in a strictly decreasing manner. Let us insert such a sequence $\underline{\alpha}$. Since h

---

[6]For an introduction and comprehensive survey of nonstandard analysis, cf. e.g. Albeverio, Fenstad, Høegh-Krohn and Lindstrøm [**1**].

[7]The existence of non-principal ultrafilters is a consequence of, though weaker than, the Axiom of Choice, cf. Halpern and Levy [**19**]

is a linear weighted averaging operation and all entries of $\underline{\alpha}$ are positive, $h(\underline{\alpha}) > 0$. However, since $\underline{\alpha}$ is a null sequence, $[\underline{\alpha}]_{\sim_{\mathcal{U}}}$ is an infinitesimal on account of Łoś's theorem, hence $^\circ[\underline{\alpha}]_{\sim_{\mathcal{U}}} = 0$, contradicting Equation (6).

We have to prove that F need not be continuous linear. Now, if N is countable, then $\mathcal{U}$ cannot be σ-complete, and therefore the set function $\chi_{\mathcal{U}}$ defined on 2^N is not a σ-additive measure, whence the aggregation functional F does not admit a classical integral representation (and hence is not continuous linear).

If N is uncountable, then $\mathcal{U}$ can at the very least be chosen in such a manner that $\mathcal{U}$ is not σ-complete and hence F is not continuous linear: For example, suppose that either there are no weakly inaccessible cardinals or the cardinality of N is less than the least such cardinal. If $\mathcal{U}$ were σ-complete on N, then it would induce a σ-additive measure $\chi_{\mathcal{U}}$ on N, which contradicts our cardinality assumption (cf. Ulam [52], Jech [28, p. 126, Theorem 10.1]). Hence, unless the cardinality of N is extremely large, no non-principal ultrafilter $\mathcal{U}$ can be σ-complete, whence F will not be continuous linear. (Note, that the existence of weakly inaccessible cardinals cannot be proved from **ZFC**, cf. Jech [28, p. 33].) In general, regardless of the cardinality of N, there will at least exist non-principal ultrafilters $\mathcal{U}$ which are not σ-complete (and satisfy additional properties, cf. e.g. Kunen [34, Theorem 3.2]).

Finally, we have to verify that F is finitely anonymous. To see this, note that if $\pi : N \to N$ is a permutation which is constant on a co-finite subset of N, then $I_\pi = \{i \in N : \pi(i) = i\} \in \mathcal{U}$ by our choice of $\mathcal{U}$. As $\mathcal{U}$ is closed under supersets, this means that for all $\underline{\beta} \in [0,1]^N$, $\underline{\beta} \sim_{\mathcal{U}} (\beta_{\pi(i)})_{i \in N}$, i.e.,

$$(7) \qquad [\underline{\beta}]_{\sim_{\mathcal{U}}} = \left[(\beta_{\pi(i)})_{i \in N} \right]_{\sim_{\mathcal{U}}}..$$

Now let $\mathcal{A} \in \Sigma$, $\underline{P} \in \Delta(\mathcal{A})^N$ and $A \in \mathcal{A}$, and put $\underline{P}' = (P_{\pi(i)})_{i \in N}$. Then,

$$[\underline{P}(A)]_{\sim_{\mathcal{U}}} = [\underline{P}'(A)]_{\sim_{\mathcal{U}}}$$

by Equation (7), hence $F(\underline{P})(A) = F(\underline{P}')(A)$ by definition of F. $\qquad \square$

PROOF OF THEOREM 10. • The equivalence follows from Lemma 11. If the vector $\underline{\alpha}$ in that Lemma is chosen to contain only non-zero entries, the resulting aggregation functional is clearly non-oligarchic

- If N is countably infinite, there are of course continuous linear aggregation functionals, and whenever the measure that represents such a functional has full support, the resulting aggregation functional will be non-oligarchic. There will then exist a sequence $\underline{\alpha} \in (0,1)^N$ such that $F(\underline{P})(A) = \sum_{i \in N} \alpha_i P_i(A)$ (convergent) for all $\mathcal{A} \in \Sigma$, $\underline{P} \in \Delta(\mathcal{A})^N$ and $A \in \mathcal{A}$. (In particular, $\sum_{i \in N} \alpha_i = 1$.) However, Lemma 12 teaches that there also other aggregation functionals which are non-oligarchic and systematic/SSFP.

- This follows from Lemma 12.

- The existence of a finitely anonymous continuous linear functional entails the existence of a σ-additive measure on N which is "nontrivial" in the sense of set-theoretic measure theory (cf. Jech [28, p. 125, (10.1 iii)]), i.e. it does not assign positive mass

to a singleton. This, however, implies the existence of a weakly inaccessible — in fact, even of a measurable — cardinal (cf. Ulam [**52**], Jech [**28**, p. 126, Theorem 10.1]). This, however, is not provable from **ZFC** (cf. Jech [**28**, p. 33]).[8]

- Since no successor cardinal is weakly inaccessible, the above argument shows that there cannot be a continuous linear aggregation functional if N has successor cardinality.

$\square$

6. Discussion and conclusion

Formal epistemologists who accept Knightian uncertainty (or ambiguity) — i.e. the thesis that rational agents may have belief systems that are compatible with several subjective probability measures (multiple priors) — need to come up with a new account of how such agents will make their decisions, as classical expected utility theory requires, of course, a unique prior. The standard decision-theoretic literature (e.g. Gilboa and Schmeidler [**18**]) treats Knightian uncertainty through axiomatic rationalisation of generalisations of expected-utility preferences, but this is clearly not satisfactory from an epistemological point of view, since there are situations where decisions have to be rationalised with respect to an aggregate belief system (as, for instance, in Bradley's [**3**] psychologically informed account of individual decisions, or in social epistemology, cf. Bradley, Dietrich and List [**4**]). What is needed is a natural aggregation theory for probability measures — ideally one that is applicable to infinite profiles of priors.

While a number of obvious approaches (such as Arrovian aggregation of expected-utility preferences or their generalisations) turn out to be barren, there does exist a candidate for such an aggregation theory of (possibly infinitely many) probability measures: We have proved an extension of McConway's [**43**] results on probabilistic opinion pooling, which can be regarded as Arrovian [**2**] in spirit, as it relates to Arrovian social choice theory through the unified general aggregation theory of Dietrich and List [**12, 13**]. The existence of well-behaved aggregation functionals for uncountably infinite profiles is non-trivial, since obvious candidates such as integral-based aggregators quickly lead into (set-theoretic) measurability problems. Moreover, the aggregation functionals that we construct for the case of infinite electorates also satisfy a weak anonymity condition.

Now some readers might be concerned about the use of the ultrafilter existence theorem and nonstandard analysis in the construction of the aggregation functionals for infinite electorates. It may be reassuring to recall that (i) any proof invoking nonstandard analysis can always be transformed into a long standard proof; (ii) that the ultrafilter existence theorem is strictly weaker than the Axiom of Choice (Halpern and Levy [**19**]); and (iii) that there are nonstandard models of the reals — even nonstandard universes — which are definable over Zermelo–Fraenkel set theory plus Choice (Kanovei and Shelah [**31**], Herzberg [**20, 21**]).

---

[8]If the existence of (weakly) inaccessible cardinals were provable, one could prove the consistency of **ZFC**, violating Gödel's second incompleteness theorem.

To conclude: It is possible to directly aggregate — finite and infinite — profiles of finitely additive probability measures in a way that (i) respects Arrovian-spirited responsiveness axioms, (ii) reduces, in the case of finite profiles, to the intuitive rule of linear averaging of probabilities, (iii) is — for several decision-theoretic reasons — the most natural viable approach to the aggregation of priors. Philosophically, this means that there is a rigorous sense in which one can refer (1) to the 'aggregate belief system' of a group of individuals who hold probabilistic beliefs (as in Bradley, Dietrich and List [**4**]) and (2) to the 'aggregate prior' of an agent whose belief system is compatible with several subjective probability measures (as recently suggested by Bradley [**3**]).

Acknowledgments

References

[1] S. Albeverio, R.J. Høegh-Krohn, J.E. Fenstad, and T. Lindstrøm. *Nonstandard methods in stochastic analysis and mathematical physics*, volume 122 of *Pure and Applied Mathematics*. Academic Press, Orlando, FL, 1986.

[2] K.J. Arrow. *Social choice and individual values*, volume 12 of *Cowles Commission Monographs*. Wiley, New York, 2 edition, 1963.

[3] R. Bradley. Bayesian conditioning revisited. Colloquium in Logic, Philosophy of Science and Philosophy, Munich Center for Mathematical Philosophy (MCMP/LMU), June 2012.

[4] R. Bradley, F. Dietrich, and C. List. Aggregating causal judgments. *Philosophy of Science*, 81:411–515, 2014.

[5] D.E. Campbell. Intergenerational social choice without the Pareto principle. *Journal of Economic Theory*, 50(2):414–423, 1990.

[6] S. Cerreia-Vioglio, P. Ghirardato, F. Maccheroni, M. Marinacci, and M. Siniscalchi. Rational preferences under ambiguity. *Economic Theory*, 48(2-3):341–375, 2011.

[7] R.M. Cooke. *Experts in uncertainty. Opinion and subjective probability in science.* Environmental Ethics and Science Policy Series. Oxford University Press, New York, 1991.

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

[9] B. de Finetti. *Teoria delle probabilità: sintesi introduttiva con appendice critica. Volumi primo e secondo.* Nuova Biblioteca Scientifica Einaudi. Giulio Einaudi Editore, Turin, 1970.

[10] B. de Finetti. *Theory of probability: a critical introductory treatment. Vol. 1. Translated by Antonio Machì and Adrian Smith, with a foreword by D. V. Lindley.* Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, London, New York, Sydney, 1974.

[11] B. de Finetti. *Theory of probability: a critical introductory treatment. Vol. 2. Translated from the Italian by Antonio Machì and Adrian Smith.* Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, London, New York, Sydney, 1975.

[12] F. Dietrich and C. List. Opinion pooling on general agendas. Working paper, 2008. http://franzdietrich.net/Papers/DietrichList-OpinionPooling.pdf.

[13] F. Dietrich and C. List. The aggregation of propositional attitudes: Towards a general theory. volume 3 of *Oxford Studies in Epistemology*, pages 215–234. Oxford University Press, Oxford, 2010.

[14] K. Easwaran and B. Fitelson. An "evidentialist" worry about Joyce's argument for probabilism. *Dialectica*, 66(3):425–433, 2012.

[15] D. Ellsberg. Risk, ambiguity, and the Savage axioms. *Quarterly Journal of Economics*, 75(4):643–669, 1961.

[16] M. Fey. May's theorem with an infinite population. *Social Choice and Welfare*, 23(2):275–293, 2004.

[17] B. Fitelson and D. McCarthy. Steps toward a new foundation for subjective probability. Work in progress, Munich Center for Mathematical Philosophy, 2012.

[18] I. Gilboa and D. Schmeidler. Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics*, 18(2):141–153, 1989.

[19] J.D. Halpern and A. Levy. The Boolean prime ideal theorem does not imply the axiom of choice. In D. Scott, editor, *Axiomatic Set Theory*, volume XIII, part 1 of *Proceedings of Symposia in Pure Mathematics*, pages 83–134. American Mathematical Society, Providence, RI, 1971.

[20] F.S. Herzberg. A definable nonstandard enlargement. *Mathematical Logic Quarterly*, 54(2):167–175, 2008.

[21] F.S. Herzberg. Addendum to "A definable nonstandard enlargement". *Mathematical Logic Quarterly*, 54(6):666–667, 2008.

[22] F.S. Herzberg. Aggregation of Monotonic Bernoullian Archimedean preferences: Arrovian impossibility results. IMW Working Paper 488, Center for Mathematical Economics, Bielefeld University, September 2013.

[23] F.S. Herzberg. The (im)possibility of collective risk measurement: Arrovian aggregation of variational preferences. *Economic Theory Bulletin*, 1(1):69–92, 2013.

[24] F.S. Herzberg. Universal algebra for general aggregation theory: Many-valued propositional-attitude aggregators as MV-homomorphisms. *Journal of Logic and Computation*, 25(3):965–977, 2015.

[25] F.S Herzberg and D. Eckert. Impossibility results for infinite-electorate abstract aggregation rules. *Journal of Philosophical Logic*, 41(1):273–286, 2012.

[26] F.S. Herzberg and D. Eckert. The model-theoretic approach to aggregation: Impossibility results for finite and infinite electorates. *Mathematical Social Sciences*, 64(1):41–47, 2012.

[27] A. Hylland and R.J. Zeckhauser. The impossibility of Bayesian group decision making with separate aggregation of beliefs and values. *Econometrica*, 47(6):1321–1336, 1979.

[28] Th. Jech. *Set theory. The third millennium edition.* Springer Monographs in Mathematics. Springer, Berlin, 2000.

[29] R.C. Jeffrey. *Probability and the art of judgment.* Cambridge Studies in Probability, Induction, and Decision Theory. Cambridge University Press, Cambridge, 1992.

[30] J.M. Joyce. Accuracy and coherence: Prospects for an alethic epistemology of partial belief. In F. Huber and C. Schmidt-Petri, editors, *Degrees of Belief*, volume 342 of *Synthese Library*, pages 263–297. Springer, Dordrecht, 2009.

[31] V. Kanovei and S. Shelah. A definable nonstandard model of the reals. *Journal of Symbolic Logic*, 69(1):159–164, 2004.

[32] C. Klamler and D. Eckert. A simple ultrafilter proof for an impossibility theorem in judgment aggregation. *Economics Bulletin*, 29(1):320–328, 2009.

[33] F.H. Knight. *Risk, uncertainty, and profit.* Hart, Schaffner & Marx, Boston, Massachusetts, 1921.

[34] K. Kunen. Ultrafilters and independent sets. *Transactions of the American Mathematical Society*, 172:299–306, 1972.

[35] L. Lauwers and L. Van Liedekerke. Ultraproducts and aggregation. *Journal of Mathematical Economics*, 24(3):217–237, 1995.

[36] M. Le Breton. Essais sur les fondements de l'analyse économique de l'inégalité. Thèse pour le Doctorat d'État en Sciences Économiques, Université de Rennes 1, 1986.

[37] H. Leitgeb and R. Pettigrew. An objective justification of Bayesianism I. Measuring inaccuracy. *Philosophy of Science*, 77(2):201–235, 2010.

[38] H. Leitgeb and R. Pettigrew. An objective justification of Bayesianism II. The consequences of minimizing inaccuracy. *Philosophy of Science*, 77(2):236–272, 2010.

[39] D. Lindley. Reconciliation of probability distributions. *Operations Research*, 31(5):866–880, 1983.

[40] P.A. Loeb. Conversion from nonstandard to standard measure spaces and applications in probability theory. *Transactions of the American Mathematical Society*, 211:113–122, 1975.

[41] J. Łoś. Quelques remarques, théorèmes et problèmes sur les classes définissables d'algèbres. In Th. Skolem, G. Hasenjaeger, G. Kreisel, A. Robinson, H. Wang, L. Henkin, and J. Łoś, editors, *Mathematical interpretation of formal systems*, volume 16 of *Studies in Logic and the Foundations of Mathematics*, pages 98–113. North-Holland, Amsterdam, 1955.

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

[43] K.J. McConway. Marginalization and linear opinion pools. *Journal of the American Statistical Association*, 76(374):410–414, 1981.

[44] M. Minsky. *The society of mind*. Simon & Schuster, New York, 1986.

[45] R.E. Ornstein. *Multimind*. Macmillan, London, 1986.

[46] F. Riedel. Optimal stopping with multiple priors. *Econometrica*, 77(3):857–908, 2009.

[47] A. Robinson. Non-standard analysis. *Nederlandse Akademie van Wetenschappen. Proceedings. Series A. Indagationes Mathematicae*, 64:432–440, 1961.

[48] A. Robinson. *Non-standard analysis*. North-Holland, Amsterdam, 1966.

[49] L.J. Savage. *The foundations of statistics*. John Wiley & Sons, New York, 1954.

[50] R.C. Schwartz. *Internal Family Systems Therapy*. Guilford Press, New York, 1997.

[51] M.H. Stone. The theory of representations for Boolean algebras. *Transactions of the American Mathematical Society*, 40(1):37–111, 1936.

[52] S. Ulam. Zur Maßtheorie in der allgemeinen Mengenlehre. *Fundamenta Mathematicae*, 16:140–150, 1930.

[53] J. von Neumann and O. Morgenstern. *Theory of games and economic behavior*. Princeton University Press, Princeton, New Jersey, 1944.

[54] R. Wedgwood. Outright belief. *Dialectica*, 66(3):309–329, 2012.