Bruno de Finetti and Imprecision: Imprecise Probability Does not Exist!

Paolo Vicig[*,a], Teddy Seidenfeld[b]

[a]*University of Trieste, Italy*
[b]*Carnegie Mellon University, USA*

Abstract

We review several of de Finetti's fundamental contributions where these have played and continue to play an important role in the development of imprecise probability research. Also, we discuss de Finetti's few, but mostly critical remarks about the prospects for a theory of imprecise probabilities, given the limited development of imprecise probability theory as that was known to him.

Key words: Coherent previsions, imprecise probabilities, indeterminate probabilities

1. Introduction

Researchers approaching the theory of imprecise probabilities [IP] may easily deduce that Bruno de Finetti's ideas were influential for its development.

Consider de Finetti's foundational *Foresight* paper (1937), which is rightly included in the first volume of the series *Breakthroughs in Statistics* [18]. In that paper we find fundamental contributions to the now familiar concepts of *coherence* of subjective probabilities – having fair odds that avoid sure loss – and *exchangeable* random variables – where permutation symmetric subjective probabilities over a sequence of variables may be represented by mixtures of *iid* statistical probabilities. Each of these concepts is part of

---

[*]Corresponding author. Tel.: +39 040 5587108; fax: +39 040 5587069.
   Email addresses: `paolo.vicig@econ.units.it` (Paolo Vicig), `teddy@stat.cmu.edu` (Teddy Seidenfeld)

the active research agendas of many IP researchers. That is, we continue to see advances in IP that are based on novel refinements of coherence, and contributions to concepts of probabilistic independence as those relate also to *exchangeability*. For instance, 7 of 47 papers in the *Proceedings* (2009) of the 6th *ISIPTA - International Symposium on Imprecise Probability: Theories and Applications* include at least one citation of de Finetti's work. And it is not hard to argue that another 7, at least, rely implicitly on his fundamental contributions.

Regarding the origins of modern studies on IP, consider for instance Walley's book [48], nowadays probably the best known extensive treaty on imprecise probabilities. Key concepts like upper and lower previsions, their behavioural interpretation, the consistency notions of coherence and of previsions that avoid sure loss, appear at once as generalizations of basic ideas from de Finetti's theory. In the preface to [48], Walley acknowledges that

> '*My view of probabilistic reasoning has been especially influenced by the writings of Terrence Fine, Bruno de Finetti, Jack Good, J.M. Keynes, Glenn Shafer, Cedric Smith and Peter Williams*'.

In their turn, most of these authors knew de Finetti's theory, while Smith [41] and especially Williams [51] were largely inspired by it. From this, and noting also the many places in Walley's book discussing de Finetti's ideas (see especially Sects. 2.1.4, 2.3.6, 5.7.1, 5.7.6, 6.8.4, 6.9.7, 9.5.6 in [48]), one may easily infer the outstanding relevance of de Finetti's thought in Walley's approach.

For another intellectual branch that has roots in de Finetti's work, consider contributions to IP from Philosophy. For example, Levi [28, 29] generalizes de Finetti's decision–theoretic concept of *coherence* through his rule of *E–admissibility* applied with convex sets of credal probabilities and cardinal utilities.

However, a closer look at de Finetti's writings demonstrates that imprecise probabilities were a secondary issue in his work, at best. He did not write very much about them. In fact, he was rather skeptical about developing a theory based on what he understood IP to be about. To understand the incongruity between the incontrovertible fact that many IP researchers recognize the origins for their work in de Finetti's ideas but that de Finetti did not think there was much of a future in IP, we must take into account the historical context in the first half of the last century, and the essentially

marginal role in the scientific community of the few papers known at the time that treated imprecision by means of alternatives to precise probability.

Our note is organized as follows: In Section 2 we discuss de Finetti's viewpoint on imprecision. After reviewing some historical hints (Section 2.1), we summarize what we understand were de Finetti's thoughts on IP (Section 2.2). In Section 3 we respond to some of de Finetti's concerns about IP from the current perspective, i.e., using arguments and results that are well known now but were not so at the earlier time. We review some key aspects of the influence of de Finetti's thought in IP studies in Section 4. Section 5 concludes the paper.

2. Imprecise Probabilities in de Finetti's Theory

2.1. A Short Historical Note

De Finetti published his writings over the years 1926–1983, and developed a large part of his approach to probability theory in the first thirty years. In the first decade (1926–1936) he wrote about seventy papers, the majority on probability theory. At the beginning of his activity, measure–theoretic probability was a relatively recent discipline attracting a growing number of researchers. There was much interest in grounding probability theory and its laws (Kolmogorov's influential and measure–theoretic approach to probability was published in 1933), and few thought of other ways of quantifying uncertainty. Yet, alternatives to probability had already been explored: even in 1713, more or less at the origins of probability as a science, J. Bernoulli considered non-additive probabilities in Part IV of his *Ars Conjectandi*, but this aspect of his work was essentially ignored (with the exception of J.H. Lambert, who derived a special case of Dempster's rule in 1764 ([37], p. 76)). Other questions which play an important role in de Finetti's work, but not in the standard measure–theoretic approach, had been investigated by G. Boole [2]; see especially the logical rather than set–theoretic interpretation of events and the analysis of the probability bounds induced on additional events by a probability assessment.

In the time between Bernoulli's work and the sixties of last century, some researchers were occasionally concerned with imprecise probability evaluations, but generally as a collateral problem in their approaches. Among them, de Finetti quotes (in [16], p. 133, and [17]) B.O. Koopman and I.J. Good, asserting that the introduction of numerical values for upper and lower probabilities was a specific follow–up of older ideas by J.M. Keynes [26].

Starting from the sixties, works focusing on various kinds of imprecise probabilities appeared with slowly increasing frequency. Their authors originally explored different areas, including non-additive measures (Choquet, whose monograph [3] remained virtually unknown when published in 1954 and was rediscovered several years later), Statistics [9], Philosophy [27, 28, 42, 47], robustness in statistics [24, 25], belief functions [37]. See e.g. [22] for a recent historical note.

Among these, de Finetti certainly read two papers which referred to his own approach, [41] and [51]. While Smith's paper [41] was still a transition work, Williams' [51] technical report stated a new, in-depth theory of imprecise conditional previsions, which generalized de Finetti's betting scheme to a conditional environment, proving important results like the envelope theorem. Consider that at the time (1975) the theory of conditional (precise) previsions was not fully developed yet, and never was by de Finetti: papers on this appeared only some ten years later [23, 35]. Thus Williams produced ideas that could be, and actually are from today's perspective, taken as very broad generalizations of de Finetti's theory.

De Finetti's reaction to Smith's paper was essentially negative and, as he explained, led to the addition of two short sections in the final version of [16]. We discuss de Finetti's reactions below.

As for Williams' paper, de Finetti read it in a later phase of his activity, the mid-seventies, and we are aware of no written comments on it. However Williams commented on this very point many years later, in an interview published in *The SIPTA Newsletter*, vol. 4 (1), June 2006. In his words:

> *De Finetti himself thought the 1975 paper was too closely connected to "formal logic" for his liking, which puzzled me, though he had expressed interest and pleasure in the earlier 1974 paper linking subjective probability to the idea of the indeterminacy of empirical concepts.*

Throughout his career de Finetti proposed original ideas that were often out of the mainstream. For example, he championed the use of finite additivity as opposed to the more restrictive, received theory of countably additive probability, both regarding unconditional and conditional probability. Criticism from the prevailing measure theoretic approach to probability often dubbed finitely additive subjective probability as arbitrary. It might have been too hard to spread the even more innovative concepts of imprecise probabilities.

This may be a motivation for de Finetti's caution towards imprecise probabilities. It certainly contributes to our understanding why Williams' report [51] was published [52] only in 2007, more than thirty years later. (See [46].)

2.2. Imprecision in de Finetti's Papers

In very few places in his large body of written work does de Finetti discuss imprecise probabilities, and nowhere does he do so exclusively. Discussions of some length appear in [14, 16, 17]. De Finetti's basic ideas on imprecision appear already in the philosophical, qualitative essay [14] *Probabilismo. Saggio critico sulla teoria delle probabilità e sul valore della scienza*, which de Finetti quotes in his autobiography in [19] as the first description of his viewpoint on probability. In this paper, he acknowledges that an agent's opinion on several events is often determined up to a very rough degree of approximation, but observes that the same difficulty arises in all practical problems of measuring quantities (p. 40). He then states (p. 41) that under this perspective probability theory is actually perfectly analogous to any experimental science:

> *In experimental sciences, the world of feelings is replaced by a fictitious world where quantities have an exactly measurable value; in probability theory, I replace my vague, elusive mood with that of a fictitious agent with no uncertainty in grading the degrees of his beliefs.*

Continuing the analogy, shortly after (p. 43) he points out a disadvantage of probability theory, that

> *measuring a psychological feeling is a much more vaguely determined problem than measuring any physical quantity,*

noting however that just a few grades of uncertainty might suffice in many instances. On the other hand, he observes that the rules of probability are intrinsically precise, which allows us to evaluate the probability of various further events without adding imprecision.

In an example (p. 43, 44, abridged here), he notes that $P(A \wedge B) = P(A|B)P(B)$ is determined precisely for an agent once $P(A|B)$ and $P(B)$ are specified. By contrast, when starting from approximate evaluations like $P(B) \in [0.80, 0.95]$ and $P(A|B) \in [0.25, 0.40]$, imprecision propagates. Then $P(A \wedge B)$ can only be said to lie in the interval $[0.80 \cdot 0.25 = 0.20, 0.95 \cdot 0.40 = 0.38]$.

If B is the event: the doctor visits an ill patient at home, and A: the doctor is able to heal the ill patient, approximate evaluations – he notes – are of little use, as they do not let us conclude much more than the following merely qualitative deduction, which we paraphrase: If it is nearly sure that the doctor will come, and fairly dubious that he can heal his patient, then it is slightly more dubious that the doctor comes and heals his patient.

Further, de Finetti notes that probabilities can often be derived from mere *qualitative opinions*. For instance, in many games the atoms of a finite partition are believed to be equally likely. This remark suggests a reflection on the role of qualitative uncertainty judgements in de Finetti's work. Interestingly, he displayed a different attitude towards this definitely more imprecise tool than to imprecise probabilities. In fact, in the same year 1931 he wrote *Sul significato soggettivo della probabilità* [15], discussing rationality conditions, later known as *de Finetti's conditions*, for comparative (or qualitative) probabilities, showing their analogy with the laws of numerical probability. This paper pointed out what became an important research topic, concerning existence of agreeing or almost agreeing probabilities for comparative probability orderings. (See [21] for an excellent review.)

The ideas expressed in [14] were not substantially modified in later writings. For instance, in [16], p. 95, de Finetti and Savage quote E. Borel as sharing their thesis, that

> *the vagueness seemingly intrinsic in certain probability assessments should not be regarded as something qualitatively different from uncertainty in any quantities, numbers and data one works with in applied mathematics.*

The jointly authored 1962 paper [16], *Sul modo di scegliere le probabilità iniziali*, adds some arguments to de Finetti's ideas on imprecise probabilities while discussing Smith's then recently published paper [41]. Recall that Smith proposed a modification of de Finetti's betting scheme, introducing a one–sided lower probability $\underline{P}(A)$ and a one–sided upper probability, $\overline{P}(A) \geq \underline{P}(A)$, for an event A, rather than a single two–sided probability, as we explain next. In Smith's approach, the agent judges a bet on A (winning 1 if and only if A obtains) at a price $p < \underline{P}(A)$ to be favorable over the status quo, which has 0 payoff for sure. Such a favorable gamble has a positive lower expected value, hence greater than 0. And for the same reason the agent prefers to bet against A (paying 1 if and only if A obtains) in order to receive a price $p > \overline{P}(A)$ over the status quo. For prices p between the lower and upper

probability, $\underline{P}(A) \leq p \leq \overline{P}(A)$, the agent is allowed to abstain from betting and remain with the status quo.

In de Finetti's theory, by contrast, the agent is obliged to give one two-sided probability $P(A)$ for betting on/against the event A. At the *fair* price $p = P(A)$ the de–Finetti–agent is indifferent between a gamble on/against A and abstaining, and may either accept or reject the bet. For prices $p < P(A)$ the de–Finetti–agent judges a bet on A favorable, etc. Thus, de Finetti's theory is the special case of Smith's theory when $P(A) = \overline{P}(A) = \underline{P}(A)$, modulo the interpretation of how the agent may respond to the case of a *fair* bet.

After expressing perplexity about the idea of avoiding stating *one* exact fair value $P(A)$ by introducing an indecision interval $I = [\underline{P}(A), \overline{P}(A)]$, with *two* different exact (one-sided) values as endpoints, de Finetti and Savage focus on two questions: *first*, existence of the indecision interval I and *second*, consistency of the agent's betting using the interval I.

As for the first question, de Finetti and Savage agree that nobody is *actually* willing to accept all of the bets required according to the idealized version of de Finetti's coherence principle. They concede that the betting model introduced by de Finetti in order to give an operational meaning to *subjective probability* requires that an *idealized*, rational agent is obliged to have a real–valued probability $P(A)$ and, thus, to accept bets at favorable odds – betting on A for any price less than $P(A)$ and betting against A for any price greater than $P(A)$.[1] The real agent is committed to behave according to the idealized theory in *hypothetical circumstances* where he/she has reflected adequately on the problem. In other words, de Finetti's opinion, expressed on this point also in other papers, seems to be that the betting scheme should not be taken literally. Rather it is a way of defining the subjective probability concept in idealized circumstances. Hence, intervals of indecision exist in practice, but only where the real decision agent has not thought through the betting problem with the precision asked of the idealized agent.

As for the second question, de Finetti and Savage argue that, rather than allowing the indecision interval, from the perspective of coherence it may be better to employ the precise two–sided probability $P = (\overline{P} + \underline{P})/2$. They

---

[1] As recalled in [16], such agents were termed *Stat Rats* (by G.A. Barnard) in the discussion of [41].

report the following intriguing example as evidence for their view.

Example (de Finetti and Savage, 1962, p. 139). An agent may choose whether to buy or not any combination of the following 200 tickets involving varying gambles on/against event A. The first 100 tickets are offered for prices, respectively, of 1, 2,..., 100 Euros[2] and each one pays 100 Euros if event A occurs, and 0 otherwise. The remaining 100 tickets are offered, respectively, at the same prices but on the complementary event, A^c. Each of these 100 tickets pays 100 Euros if A^c occurs and 0 otherwise. If the agent assesses a two–sided personal probability for A as in de Finetti's theory, e.g., $P(A) = 0.63$, he/she will maximize expected value by buying the first 63 tickets on A with prices 1,..., 63, for a combined price $1 + 2 + \ldots + 63 = 2016$ Euros, and buying the first 37 tickets on A^c for a combined price $1 + 2 + \ldots + 37 = 703$ Euros. (The agent is indifferent about buying the 63rd ticket from the first group and, likewise, the 37th ticket from the second group.) The agent's total expense for the 100 tickets, then, is 2719 Euros. The agent gains $6300 - 2719 = 3581$ Euros if A occurs; he/she gains 981 Euros otherwise, when A^c occurs.

Suppose, instead the agent fixes a lower probability $\underline{P}(A) = 0.53$ and an upper probability $\overline{P}(A) = 0.73$, as allowed by Smith's theory. De Finetti and Savage interpret this to mean that the Smith-agent will buy only the first 53 tickets for A and only the first 27 tickets for A^c – those gambles that are individually (weakly) *favorable*. Then the Smith–agent will gain only $5300 - 1809 = 3491$ Euros if A occurs, and will gain only $2700 - 1809 = 891$ Euros if A^c occurs. Their conclusion is that in this decision problem it is better for the agent to assess the real–valued, two–sided probability $0.63 = P(A) = (\overline{P}(A) + \underline{P}(A))/2$ than to use the interval $I = [0.53, 0.73]$. The decision maker's gain increases by 90 Euros, whatever happens, using this two-sided, de Finetti–styled probability. We respond to this example in the next section. $\square$

De Finetti and Savage continue their criticism of IP theory on pp. 140 $\div$ 144 of [16]. To our thinking, the most interesting argument they offer is perhaps that imprecision in probability assessments does not give rise to

---

[2]We introduce an anachronism, here and in later examples, updating the monetary unit to 2012.

a new kind of uncertainty measure, but rather points out an incomplete elicitation by a third party and/or even incomplete self–knowledge. They write,

> *Even though in our opinion they are not fit for characterizing a new, weaker kind of coherent behaviour, structures and ideas like Smith's may allow for important interpretations and applications, in the sense that they elicit what can be said about a behaviour when an incomplete knowledge is available of the opinions upon which decisions are taken.*

They continue with a clarifying example.

> *What is the area of a triangle with largest side a and shortest side b? Any S such that $\underline{S} \leq S \leq \overline{S}$, with $\underline{S}$: area of the triangle with sides (a, b, b), $\overline{S}$: area of the triangle with sides (a, a, b). This does not mean: there exists a triangle whose area is indeterminate ($\underline{S}$: lower area, $\overline{S}$: upper area); every triangle has a well determined area, but we might at present be unable to determine it for lack of sufficient information.*

In the Appendix of [17], while mainly summarizing ideas on imprecise probabilities already expressed in [14, 16], de Finetti adds other examples supporting the same thesis. One is particularly interesting because it does not resort to the analogy between probabilities and other experimental measures but involves his *Fundamental Theorem of Prevision*. As well known, that theorem ensures that, given a coherent probability function $P(\cdot)$ defined on an arbitrary set of events $\mathcal{D}$, all of its coherent extensions that include a probability for an additional event $E \notin \mathcal{D}$ belong to a non-empty closed interval $I_E = [\underline{P}(E), \overline{P}(E)]$. This interval I_E of potential (coherent) values for $P(E)$ is defined by analogy with how one may extend a measure μ to give a value for a non-measurable set using the interval of inner and outer measure values. In de Finetti's theorem, the interval I_E arises by approximations to E (from below and from above) using events from the linear span of $\mathcal{D}$. But, de Finetti argues, the fact that prior to the extension, we can only affirm about $P(E)$ that it belongs to I_E rather than having a unique value

> *does not imply that some events like E have an indeterminate probability, but only that $P(E)$ is not uniquely defined by the starting data we consider.*

De Finetti's thinking about imprecise personal probability is unchanged from his early work.[3] In his classic ([36], p. 58) Savage quotes de Finetti's [18] view on this question.

> *The fact that a direct estimate of a probability is not always possible is just the reason that the logical rules of probability are useful.*

Revealing of Savage's subsequent thinking on this question of existence of unsure, or imprecise (personal) probabilities is the footnote on p. 58, added for the 1972 edition of [36], where Savage teases us with these guarded words.

> *One tempting representation of the unsure is to replace the person's single probability measure P by a set of such measures, especially a convex set. Some explorations of this are Dempster (1968), Good (1962), and Smith (1961).*

3. Rejoinder from the Perspective of 2011

Many of the objections raised by de Finetti (and others) towards the use of imprecise probabilities have been discussed at length elsewhere. (See especially [48], Secs. 5.7, 5.8, 5.9). Of course, some recently formulated arguments in favor of IP, e.g., some relating to group decision making [39] or IP models for frequency data [12], were not anticipated by de Finetti. Here, we offer brief comments, including responses to the challenges against IP raised in the previous section.

The first of de Finetti's arguments supporting precise rather than imprecise probabilities is roughly that – barring e.g., Quantum Mechanical issues – ordinary theoretical quantities that are the objects of experimental measurement are precise. In practice however, when the process for eliciting a precise personal probability is not sufficiently reliable, impractical, or too expensive, the use of imprecise probabilities seems appropriate. By modeling the elicitation process, e.g., by considering psychometric models of introspection, we may be able to formalize the degree of imprecision of the assessment [31]; a first, intuitive measure of imprecision is of course the difference $\overline{P}(A) - \underline{P}(A)$.

---

[3]Even in the commented collection [20] of the conferences de Finetti held at the Italian Institute for Advanced Mathematics in 1979, edited and published posthumously by Mura, the editor's Note 9, p. 43, concludes that *"De Finetti has opposed this [the imprecise probability] framework by dubbing it a useless and illusory attempt to eliminate the inevitable idealizations embedded in any mathematical scheme"*.

De Finetti hits the mark with his second observation, basically that inferences with imprecise probabilities may be highly imprecise. This is unquestionably true, but there are different levels: highly imprecise measures like possibilities and necessities typically ensure many vacuous inferences [50], while standard, less imprecise instruments are (now) available in other instances, e.g., the Choquet integral for 2–monotone measures [5], the imprecise Dirichlet model [49], etc.. Moreover, specific independence or irrelevance assumptions, as well as other dependance hypotheses like exchangeability, can really constrain imprecise probability values, see among others [4, 6, 7, 44].

De Finetti and Savage's [16] example, which we summarized in Section 2.2, merits several responses. First, it is not clear what general claim they make. Are they suggesting that a decision maker who uses Smith's lower and upper IP betting odds *always* makes inferior decisions compared with some de Finetti–styled decision maker who uses precise betting odds but has no other advantage – no other special information? Is their claim instead that *sometimes* the IP decisions will be inferior? What is their objection?

De Finetti and Savage's example uses particular values for $P, \overline{P}$, and $\underline{P}$, combined with a controversial (we think unacceptable) interpretation of how the IP decision maker chooses in their decision problem. It is not difficult to check that the same conclusion they reach may be achieved by varying the three quantities $P, \overline{P}$, and $\underline{P}$ subject to the constraint that $\underline{P} < P < \overline{P}$ and these belong to the set $\{0, 1/100, 2/100, \ldots, 1\}$ while retaining the same ticket prices, and the same seemingly myopic decision rule for determining which tickets the IP decision maker purchases. That is, it appears to us that what drives de Finetti and Savage's result in this example is the tacit use of a decision rule that is invalid with sets of probabilities but which is valid in the special case of precise probabilities.

We think they interpret Smith's lower and upper betting odds to mean that when offered a bet on or against an event A at a price between its lower and upper values, the IP decision maker will reject that option *regardless what other (non-exclusive) options are available.* That is, we think they reason that, because at odds between the lower and upper probabilities it is not favorable to bet either way on A *compared with the one option to abstain,* therefore the IP decision maker will abstain, i.e. not buy such a ticket in their decision problem.

The familiar decision rule to reject as inadmissible any option that fails to maximize expected utility reduces to pairwise comparisons between pairs of acts when the agent uses a precise probability. That is, in the example under

discussion where utility is presumed to be linear in the *numeraire* used for the gambles[4], a de Finetti–styled decision maker will maximize expected utility by buying each ticket that, *by itself*, has positive expected value: Buy each ticket that in a pairwise comparison with abstaining is a favorable gamble and only those. But this rule is not correct for a decision maker who uses sets of probabilities. De Finetti and Savage's conclusion about which tickets the IP decision maker will buy is incorrect when she/he uses an appropriate decision rule.

As well known, there is continuing debate about decision rules for use with an IP theory. However, for the case at hand, we think it is non-controversial that the IP decision maker will judge inadmissible any combination of tickets that is *simply dominated* in payoff by some other combination of tickets. That is, in the spirit of de Finetti's coherence condition, particularly as he formulates it with Brier score, the decision maker will not choose an option when there is a second option available that simply dominates the first. Then, in this example, it is permissible for such an IP decision maker to buy the very same combination of tickets as would any de Finetti–styled decision maker who has a precise personal probability for the event A. That is because, in this finite decision problem, all and only Bayes–admissible options are undominated. Thus, it is impermissible for the IP decision maker to buy only the $80 = (53 + 27)$ tickets that de Finetti and Savage allege will be purchased.

Call *House* the vendor of the 200 tickets. *House* is clearly incoherent. In fact, an agent can make arbitrage without needing to consider her/his uncertainty about the event A: buying the first 50 tickets for A and the first 50 for A^c produces a sure gain of 2450 Euros! See [40] for different indices for the degree of incoherence displayed by *House*, what strategies maximize the sure gains that can be achieved against *House*, and how these are related to different IP models for the events in question.

There is a related point about IP-coherence that we think is worth emphasizing. Consider making a single bet in favor of A. If the decision maker adopts a precise probability $P(A)$, her/his gain per Euro staked on a bet on A will be $G = A - P(A)$. However, if the decision maker's judgment is unsure

---

[4]Linearity of utility is no real restriction, because coherence is equivalent to *constrained coherence*, where an arbitrary upper bound $k > 0$ is set *a priori* on the agent's gains/losses in absolute value (see [34], Sec. 3.4). Just choose k such that the utility variation is to a good approximation linear.

and she/he uses Smith's lower betting odds with $\underline{P}(A) < P(A)$, her/his gain increases to $\underline{G} = A - \underline{P}(A) > G$. It is true that in this latter case the decision maker will abstain from betting when the price for A is higher than $\underline{P}$ and lower than P, *and provided there are no other options to consider*. But this results only in the loss of some additional *opportunities* for gambling. There is no loss of a *sure gain*.

The role of the Fundamental Theorem in relation to IP theory is also worth discussing. Let us accept de Finetti's interpretation of the interval I_E as giving all coherent extensions of the decision maker's current probability $P(\cdot)$, defined with respect to events in the set $\mathcal{D}$, in order to include the new event E. Suppose, however, that we consider extending P to include a second additional event F as well. To use the Fundamental Theorem to evaluate probability extensions for both E and F we must work step–by–step. Extend $P(\cdot)$ to include only one of the two events E or F using either interval I_E or I_F defined with respect to the set $\mathcal{D}$. For instance, first extend P to include a precise value for $P(E)$ taken from I_E. Denote the resulting probability $P^E(\cdot)$ defined with respect to the set $\mathcal{D} \cup \{E\}$. Then iterate to extend $P^E(\cdot)$ to include a precise value for $P^E(F)$. Of course, the two intervals I_F and I_F^E usually are not the same. We state without demonstration that, nonetheless, if the step–by–step method allows choosing the two values $P(E) = c$ and $P^E(F) = d$, then it is possible to reverse the steps to achieve the same pair, $P(F) = d$ and $P^F(E) = c$. Then the order of extensions is innocuous.

If instead we interpret the starting coherent probability P (defined on the linear span of $\mathcal{D}$) as a special coherent lower probability, and look for a lower probability which coherently extends it, we can *avoid the step–by–step procedure*, simply by always choosing the lower endpoint from the intervals based on the common set $\mathcal{D}$ and use their values as 1-sided lower probabilities (in the two–event example, choosing the lower endpoints of the intervals I_E and I_F as lower probability assessments for, respectively, E and F). We obtain what Walley [48] calls the *natural extension* of P, interpreted as a coherent lower probability (actually, it is even n–monotone) on all additional events. The correctness of such a procedure depends also on the *transitivity* property of the natural extension.

There is a second consideration relevant to de Finetti's preferred interpretation of the interval I_E from the *Fundamental Theorem* relating to IP theory, which is particularly relevant in the light of Levi's [30] distinction between *imprecision* and *indeterminacy* of interval–valued probabilities. Levi's distinction is illustrated by Ellsberg's well known challenge [11].

In Ellsberg's puzzle [11] the decision maker faces decisions under risk and decisions under uncertainty simultaneously. The decision maker contemplates two binary choices: *Problem I* is a choice between two options labeled *1* and *2*, and *Problem II* is a choice between two options labeled *3* and *4*. The payoffs for these options are determined by the color of a randomly drawn chip from an urn known to contain only red, black, or yellow chips, cf. Table 1.

Table 1: Data for Ellsberg's puzzle

	The chip drawn from the urn is		
	RED	*BLACK*	*YELLOW*
option 1	1,000 Euros	0 Euros	0 Euros
option 2	0 Euros	1000 Euros	0 Euros
option 3	1,000 Euros	0 Euros	1,000 Euros
option 4	0 Euros	1,000 Euros	1,000 Euros

In *Problem I*, *option 1* pays off 1,000 Euros if the chip drawn is red, 0 Euros otherwise, i.e. if it is black or yellow. *Option 2* pays off 1,000 Euros if the chip drawn is black, 0 Euros otherwise, i.e, if the chip is red or yellow. In *Problem II*, *option 3* pays off 1,000 Euros if the chip drawn is either red or yellow, 0 if it is black. *Option 4* pays off 1,000 Euros if the chip drawn is black or yellow, 0 Euros if it is red. In addition, the urn is stipulated to contain exactly 1/3 red chips, with unknown proportions of black and yellow other than that their total is 2/3 the contents of the urn. Thus, under the assumptions for the problem, options *1* and *4* have determinate risk: they are just like a Savage *gamble* with determinate (personal) probabilities for their outcomes. However Ellberg's conditions leave options *2* and *3* as ill–defined gambles: the personal probabilities for the payoffs are not determined.

Across many different audiences with varying levels of sophistication, the modal choices are option *1* from Problem I and option *4* from Problem II. Assuming that the agent prefers more money to less, that there is no *moral hazard* relating the decision maker's choices with the contents of the urn, and that the choices reveal the agent's preferences, there is no expected utility

14

model for the modal pattern, *1* over *2* and *4* over *3*.

Table 1 illustrates how these preferences violate Savage's [36] Sure-Thing Postulate, *P2*. Since *options 1* and *2* have the same outcome (0 Euros) if the chip draw is *yellow*, by Expected Utility considerations, an agent with precise preference has those preferences depend solely on the consequences of these two options in the first two states, when the chip drawn is either *red* or *black*. Likewise with a comparison between *options 3* and *4*: they have the same outcome (1,000 Euros) if the chip draw is *yellow*. However, *options 1* and *3* have the same payoff when the chip drawn is *red* (1,000 Euros), and also when it is *black* (0 Euros). Similarly for *options 2* and *4*. Hence, by Expected Utility considerations (and in accord with Savage's Postulate *P2*) an agent with precise preferences cannot strictly prefer *option 1* over *option 2* and strictly prefer *option 4* over *option 3*. Given that the chip drawn is not *yellow*, *options 1* and *3* are the same, and *options 2* and *4* are the same.

In a straightforward IP–de–Finetti representation of this puzzle, the decision maker has a precise probability for the events {*red, black or yellow*}: $P(red) = 1/3$, $P(black \text{ or } yellow) = 2/3$. But the agent's uncertainty about black or yellow is represented by the common intervals $I_{black} = I_{yellow} = [0, 2/3]$. Under these circumstances the agent's imprecise probabilities do not dictate the choices for either problem. However, if after reflection the agent decides for option *1* over option *2* in Problem I, then (as in the *Fundamental Theorem*) this corresponds to an extension of $P(\cdot)$ where now $P(black) < 1/3$. But then $P(yellow) > 1/3$ and option *3* has greater expected utility than option *4* relative to this probability extension. Likewise, if the agent reflects first on Problem II and decides for option *4* over option *3*, this corresponds to an extension of $P(\cdot)$ where now $P(yellow) < 1/3$. Then in Problem I option *2* has greater expected utility than option *1*.

In short, under what we understand to be de Finetti's favored interpretation of the *Fundamental Theorem*, the modal Ellsberg choices are anomalous. They cannot be justified even when the agent uses the uncertainty intervals from the *Fundamental Theorem*. Levi calls this a case of *imprecise* probability intervals. Under this interpretation the agent is committed to resolving her/his uncertainty with a coherent, precise probability.

By contrast, if the agent uses the two intervals, $I_{black} = I_{yellow} = [0, 2/3]$, to identify a set of probabilities for the two events, then relative to this set neither option in either Problem is ruled out by considerations of expected utility. That is, in Problem I, for some probabilities in the set, option *1* has greater expected utility than option *2*, and for other probabilities in the set

this inequality is reversed. Likewise with the two options in Problem II. If the non–comparability between options by expected utility is resolved through an appeal to lower expected utility, e.g., as a form of security, then in Problem I the agent chooses option *1* and in Problem II the agent chooses option *4*. This is what Levi means by saying that the decision maker's IP is an *indeterminate* (not an *imprecise*) probability. With indeterminate probability, the agent is not committed to resolving uncertainty with a precise probability prior to choice.

4. De Finetti's Theory in Imprecise Probabilities

Let us repeat a simple fact. Notwithstanding what we see as de Finetti's mostly unsupportive opinions on imprecise probabilities, in the sense of IP as that is used by many nowadays, our co-researchers in this area find it appropriate to refer to his work in the development of their own. One reason for this is that many IP researchers use aspects of de Finetti's work on personal probability which often are in conflict with the more widely received but less general, classical theory, associated with Kolmogorov's measure theoretic approach.

Take for instance de Finetti's concept of a coherent *prevision* $P(X)$ of a (bounded) random quantity X, which is a generalization of a coherent probability. That special case obtains when X is the indicator function for an event, and then a prevision is a probability.

A prevision may be viewed as a finitely additive *expectation* $E(X)$ of X. But there are non-trivial differences between de Finetti's concept of *prevision* and the more familiar concept of a mathematical expectation as that is developed within the classic measure theoretic account. In order to determine the classical expectation of a random variable X, we first have to assess a probability for the events $\{\omega : X(\omega) = x\}$, or at least assess a cumulative distribution function. In uncountable state spaces, common with familiar statistical models, the classical theory includes measurability constraints imposed by countable additivity. But this is not at all necessary for assessing a prevision, $P(X)$, which may be determined directly within de Finetti's theory free of the usual measurability constraints. The difference may seem negligible, but it becomes more appreciable when considering previsions for several random quantities at the same time, and by far more so when passing to imprecise previsions, where additivity in general no longer applies. This

is an illustration of how de Finetti's foundational ideas can become *more important* in IP theory than they are even in traditional probability theory.

The problem reiterates within the theory of conditional expectations, magnified by the fact that finitely additive conditional expectations do not have to satisfy what de Finetti called *conglomerability*, first in his 1930 paper *Sulla proprietà conglomerativa delle probabilità subordinate* [13]. Assume that $P(\cdot)$ is a coherent unconditional probability. Let $\pi = \{h_1, \ldots\}$ be a denumerable partition, and let $\{P(\cdot|h_i) : i = 1, \ldots\}$ be a set of corresponding coherent conditional probability functions for P, given each element of π. With respect to an event E, define $m_E = \inf_{h \in \pi} P(E|h)$, and $M_E = \sup_{h \in \pi} P(E|h)$. These conditional probabilities for event E are *conglomerable* in π provided that $P(E) \in [m_E, M_E]$. Schervish et al. [38] establish that each finitely but not countably additive probability fails to be conglomerable for some event E and denumerable partition π. Also, they identify the greatest lower bound for the *extent of non–conglomerability* of P, where that is defined by the supremum difference between the unconditional probability $P(E)$ and the nearest point to the interval $[m_E, M_E]$, taken over all denumerable partitions π and events E.

The treatment of conglomerability in IP is still controversial. While Walley [48] imposes some conglomerability axioms to his concepts of coherence for conditional lower previsions, Williams' more general approach does not. In Walley's words ([48], p. 644)

> *Because it [...] does not rely on the conglomerative principle, Williams' coherence is also a natural generalization of de Finetti's (1974) definition of coherence.*

See [33], Secs. 3.4, 4.2.2 for a further discussion of [13], Williams' coherence and of some arguments in favor/against conglomerativity in IP theory.

Also de Finetti's use of a generalized betting scheme to define coherent previsions serves as an example for several subsequent variants, which underly many uncertainty measures. Examples include *coherent upper and lower previsions* [51, 48], *convex previsions* [34], and *capacities* ([1], Sec. 4). Moreover, in all such instances this approach based on de Finetti's theory of previsions provides vivid, immediate interpretations of basic concepts and often relatively simple proofs of important results.

Another issue, which was our focus in the previous section, concerns de Finetti's attention to extension problems, i.e. to the existence of at least one coherent extension of a coherent prevision, defined on an arbitrary set

of (bounded) variables. Walley [48] used this idea in the realm of imprecise probabilities to define several useful notions: a *natural* extension; a *regular* extension; an *independent* extension, etc. For instance, a natural extension is the largest, i.e., "least committal" coherent IP extension.

In general, research in IP theory exposes new facets of probability concepts already discussed and sometimes not quite fixed by de Finetti. An illustration is with the notion of *stochastic independence*, which de Finetti found unconvincing in its classical identification with the factorization property, but which he left somewhat undeveloped in his own work. In [17] he gives an epistemically puzzling example of two random quantities that are functionally dependent *and* stochastically independent according to the factorization property. Problems for a theory of independence arise especially when conditioning on events of extreme (0 or 1) probability. For instance, Dubins' version [10] of de Finetti's theory leads to an asymmetric relevance relation. The situation is more complex in the IP framework, and de Finetti would perhaps be surprised at the variety of independence concepts that have been developed. (See, e.g., [7, 8, 43, 45]).

De Finetti discovered important connections between *independence* and *exchangeability* as reported in his *Representation Theorem*, 1937. IP generalizations are being developed, e.g., [6]. Soon, we will likely see IP generalizations of *partial exchangeability* along the same lines. In yet other settings, IP methods have been employed to achieve advances in probability problems to which de Finetti himself contributed [32].

5. Conclusions

Probability does not exist is one of de Finetti's most striking mottoes, reported also on the memorial tablet at his birthplace in Innsbruck, Austria. We may possibly interpret it as a statement that probability is not something intrinsic in the laws of nature, but rather a concept of the human mind, developed for reasoning under uncertainty. In de Finetti's view imprecise probability seems not to exist in a stronger sense, not even as a really worthy tool for studying uncertainty. Yet the recent developments do not quite confirm his attitude. Thinking of this, we close our comments with this metaphor, which will be entirely familiar to any parent. You raise your children with an eye for the day when each becomes an independent agent. Sometimes, however, contrary to your advice, one embarks on what you fear

is an ill conceived plan. When to your great surprise the plan succeeds, does not that offspring then make you a very proud parent?!

Acknowledgements

References

[1] P. Baroni, R. Pelessoni and P. Vicig. Generalizing Dutch risk measures through imprecise previsions. *Internat. J. Uncertain. Fuzziness Knowledge-Based Systems*, 17:153–177, 2009.

[2] G. Boole. *An Investigation of the Laws of Thought, on which are Founded the Mathematical Theories of Logic and Probabilities.* Macmillan, 1854. Reprinted in 1958 by Dover, NY.

[3] G. Choquet. Theory of capacities. *Ann. Inst. Fourier* 5:131–295, 1954.

[4] G. de Cooman, F. Hermans, A. Antonucci, M. Zaffalon. Epistemic irrelevance in credal networks: the case of imprecise Markov trees. *Proceedings of ISIPTA'09*, 149–158, 2009.

[5] G. de Cooman, M. C. M. Troffaes and E. Miranda. n-Monotone lower previsions. *Journal of Intelligent & Fuzzy Systems*, 16: 253–263, 2005.

[6] G. de Cooman, E. Quaeghebeur and E. Miranda. Exchangeable lower previsions. *Bernoulli* 15:721–735, 2009.

[7] F. G. Cozman and P. Walley. Graphoid properties of epistemic irrelevance and independence. *Annals of Mathematics and Artificial Intelligence* 45:173–195, 2005.

[8] I. Couso, S. Moral and P. Walley. A survey of concepts of independence for imprecise probabilities. *Risk Decision and Policy* 5:165–181, 2000.

[9] A.P. Dempster. Upper and lower probabilities induced by a multivalued mapping. *Ann. Math. Stat.* 38:325–339, 1967.

[10] L.E. Dubins. Finitely additive conditional probabilities, conglomerability and disintegrations. *Ann. Prob.* 3:89–99, 1975.

[11] D. Ellsberg. Risk, ambiguity, and the Savage axioms. *Quarterly J. of Economics* 75:643–669, 1961.

[12] P.I. Fierens and T.L. Fine. Towards a frequentist interpretation of sets of measures. *Proceedings of ISIPTA'01*, 179–187, 2001.

[13] B. de Finetti. Sulla proprietà conglomerativa delle probabilità subordinate. *Rendiconti R. Istituto Lombardo di Scienze e Lettere* 43:339–343, 1930.

[14] B. de Finetti. *Probabilismo. Saggio critico sulla teoria delle probabilità e sul valore della scienza.* In: *Biblioteca di filosofia*, 1–57. Libreria editrice Perrella, Napoli, 1931.

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

[16] B. de Finetti and L.J. Savage. Sul modo di scegliere le probabilità iniziali. *Biblioteca del Metron*, Ser. C, Vol. I:81–154, 1962.

[17] B. de Finetti. *Teoria delle probabilità.* Einaudi, Torino, 1970 (English translation: B. de Finetti. *Theory of Probability*, Wiley, 1974).

[18] B. de Finetti [1937] *Foresight: Its Logical Laws, Its Subjective Sources.* (H.E. Kyburg, Jr. translation.) In: *Breakthroughs in Statistics* (S. Kotz and N.J. Johnson eds.), 1:134-174, Springer–Verlag, N.Y., 1993.

[19] B. de Finetti. *Scritti (1926–1930).* Cedam, Padova, 1981.

[20] B. de Finetti. *Philosophical Lectures on Probability*. Springer Verlag, 2008.

[21] P. Fishburn. The axioms of subjective probability (with discussion). *Stat. Sci.* 1:333–358, 1984.

[22] F. Hampel. *Nonadditive Probabilities in Statistics*. In: *Imprecision in Statistical Theory and Practice* (P. Coolen–Schrijner, F. Coolen, M. Troffaes, T. Augustin and S. Gupta eds.), 13–25, Grace Scientific Publ., Greenboro, USA, 2009.

[23] S. Holzer. On coherence and conditional prevision. *J. Boll. Un. Mat. Ital. C (6)*, Serie VI, 4(1):441–460, 1985.

[24] P.J. Huber, V. Strassen. Minimax tests and the Neyman–Pearson lemma for capacities. *Annals of Statistics* 1:251–263, 1973.

[25] P.J. Huber. *Robust Statistics*. Wiley, N.Y., 1981.

[26] J.M. Keynes. *A Treatise on Probability*. Macmillan and Co., London, 1921.

[27] H.E. Kyburg, Jr. *Probability and the Logic of Rational Belief*. Middletown, Conn., 1961.

[28] I. Levi. On indeterminate probabilities. *J.Phil.* 71:391–418, 1974.

[29] I. Levi. *Enterprise of Knowledge*. MIT Press, Cambridge, MA, 1980.

[30] I. Levi. Imprecise and indeterminate probabilities. *Proceedings of ISIPTA'99*, 258–265, 1999.

[31] D.V. Lindley, A. Tversky, and R.V. Brown. On the reconciliation of probability assessments (with discussion). *J. Roy. Statist. Soc.* Ser. A, 142:146–180, 1979.

[32] E. Miranda, G. de Cooman and E. Quaeghebeur. Finitely additive extensions of distribution functions and moment sequences: The coherent lower prevision approach. *Int. J. Approx. Reasoning* 48:132–155, 2008.

[33] R. Pelessoni and P. Vicig. Williams coherence and beyond. *Int. J. Approx. Reasoning* 50(4):612–626, 2009.

[34] R. Pelessoni and P. Vicig. Uncertainty modelling and conditioning with convex imprecise previsions. *Int. J. Approx. Reasoning* 39(2–3):297–319, 2005.

[35] E. Regazzini. Coherent conditional probabilities, finite additivity, extensions. Technical report, Dipartimento di Scienze Statistiche, University of Bologna, 1983.

[36] L.J. Savage. *The Foundations of Statistics*. John Wiley, New York, 1954.

[37] G. Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, Princeton, NY, 1976).

[38] M.J. Schervish, T. Seidenfeld, and J.B. Kadane. The extent of non–conglomerability of finitely additive probabilities. *Z.f. Wahrschein-lichkeitstheorie* 66:205–226, 1984.

[39] T. Seidenfeld, J.B. Kadane, and M.J. Schervish. On the shared preferences of two bayesian decision makers. *J.Phil.* 86:225–244, 1989.

[40] M.J. Schervish, T. Seidenfeld, and J.B. Kadane. How sets of coherent probabilities may serve as model for degrees of incoherence. *Proceedings of ISIPTA'99*, 319–323, 1999.

[41] C.A.B. Smith. Consistency in statistical inference and decision (with discussion) *J. Roy. Statist. Soc.* Ser. B, 23:1–37, 1961.

[42] P. Suppes. The measurement of belief. *J. Roy. Statist. Soc.* Ser. B 36:160-175, 1974.

[43] B. Vantaggi. Graphical representation of asymmetric graphoid structures. *Proceedings of ISIPTA'03*, 560–574, 2003.

[44] J. Vejnarova. On conditional independence in evidence theory. *Proceedings of ISIPTA'09*, 431–440, 2009.

[45] P. Vicig. Epistemic independence for imprecise probabilities. *Int. J. Approx. Reasoning* 24:235–250, 2000.

[46] P. Vicig, M. Zaffalon, and F.G. Cozman. Notes on "Notes on conditional previsions". *Int. J. Approx. Reasoning* 44:358–365, 2007.

[47] P. Walley and T. Fine. Varieties of modal (classificatory) and comparative probability. *Synthese* 41:321–374. 1979.

[48] P. Walley. *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, 1991.

[49] P. Walley. Inferences from multinomial data: learning about a bag of marmles. *J. Roy. Statist. Soc.* Ser. B 58:3–57, 1996.

[50] P. Walley and G. de Cooman. Coherence of rules for defining conditional possibility. *Int. J. Approx. Reasoning* 21 (1):63–107, 1999.

[51] P. M. Williams. Notes on conditional previsions. *Research Report*, School of Math. and Phys. Science, University of Sussex, 1975.

[52] P. M. Williams. Notes on conditional previsions. (Revised version of [51]). *Int. J. Approx. Reasoning* 44:366–383, 2007.