A BEHAVIOURAL MODEL FOR VAGUE PROBABILITY ASSESSMENTS

GERT DE COOMAN

ABSTRACT. I present an hierarchical uncertainty model that is able to represent vague probability assessments, and to make inferences based on them. This model can be given an interpretation in terms of the behaviour of a modeller in the face of uncertainty, and is based on Walley's theory of imprecise probabilities. It is formally closely related to Zadeh's fuzzy probabilities, but it has a different interpretation, and a different calculus. Through rationality (coherence) arguments, the hierarchical model is shown to lead to an imprecise first-order uncertainty model that can be used in decision making, and as a prior in statistical reasoning.

1. INTRODUCTION

It happens fairly often that people summarise their beliefs by vague linguistic statements involving probabilities, such as "the probability that the sun won't rise tomorrow is very low", or "the chances for this coin to fall heads on the next toss are close to even", or "this valve is highly unreliable". In this paper, I start from the premise that it would be interesting and useful to be able to represent this type of probability assessments, and to reason with them; and I develop a mathematical model for doing so. I have called such a model a *possibilistic prevision*, to emphasise that it constitutes a generalisation of de Finetti's previsions (or fair prices) [20] and Walley's lower and upper previsions [41], and to indicate that possibility distributions [5, 25, 50] are used to model the information conveyed by the linguistic probability assessments.

The model has a clear interpretation in terms of the behaviour of a modeller faced with uncertainty, and it therefore follows an approach to modelling uncertainty that was pioneered by Ramsey [39] and further developed by de Finetti [18], Williams [45] and Walley [41]. In spirit, it follows the same normative approach to uncertainty modelling, in that it requires that the modeller's behaviour should satisfy certain rationality, or consistency, criteria. From a mathematical point of view, it relies quite heavily on a number of basic notions and results in Walley's theory of imprecise probabilities [41].

The present model also has formal connections with Zadeh's fuzzy probabilities [51, 52], but I believe my model to be fundamentally different, since it has a clear behavioural interpretation, and a calculus that is very different from, and I shall argue less *ad hoc* than, the one suggested by Zadeh.

This paper is structured as follows. In Section 2, I give a brief review of those notions and basic results from the behavioural theory of imprecise probabilities that are necessary for understanding the material in this paper. Section 3 provides a first and more or less provisional justification for introducing possibilistic previsions. Representable possibilistic previsions are introduced in Section 4, where it is also explained why representability is a desirable consistency property for a possibilistic prevision to have. Among the representable possibilistic previsions there are special ones that I call full, and that are closely linked with the theory of fuzzy intervals. They are motivated and studied in full detail in Section 5. Section 6 explores the behavioural implications of representable possibilistic previsions, and how they can be used in decision making (and in statistical reasoning). A

Date: 20 February 2003.

Key words and phrases. fuzzy probability, probability assessments, imprecise probabilities, vagueness, second-order model, coherence, inference, decision-making.

special case of possibilistic previsions, namely those based on Walley's so-called fuzzy contamination models [42], is briefly described in Section 7. In Section 8, I supplement the provisional justification for possibilistic previsions, given in Section 3, by a more solid one, based on a behavioural interpretation in terms of so-called price functions. Section 9 concludes the paper. I have relegated a number of less central, auxiliary results to Appendices A and B.

This paper has taken me a very long time to write. The mathematical results about possibilistic previsions (without proofs) have already been published in the form of a conference paper [8] in 1998, and a fairly non-mathematical discussion of the more foundational aspects of price functions was recently published [17] as a journal paper, coauthored by Peter Walley. The present paper aims at bringing together all the relevant mathematical material in order to give a proper theoretical foundation to the notion of a possibilistic prevision, or, if you wish, a fuzzy probability.

2. BASIC NOTIONS FROM THE THEORY OF IMPRECISE PROBABILITIES

We consider a modeller who is uncertain about something, for instance, about the state of the world. We denote by Ω the set of possible states ω of the world—mutually exclusive and exhaustive—that are of interest, and call it the *possibility space*. A *gamble* on Ω is a bounded, real-valued function on the domain Ω. It is interpreted as an uncertain reward: if the true state of the world turns out to be ω then the (possibly negative) reward is $X(\omega)$. The reward X is uncertain because it is uncertain which element of Ω is the true state. For reasons of simplicity, I shall assume throughout that rewards are expressed in units of some linear utility, called *utiles*. See [41, Section 2.2] for a discussion on how to construct such a utility scale.

The set of all gambles on Ω is denoted by $\mathscr{L}(\Omega)$. It is a real linear space when provided with the point-wise addition of gambles and the point-wise multiplication of gambles with real numbers. More traditional approaches to uncertainty modelling focus on *events* rather than on gambles, but I prefer to emphasise gambles because they allow for much more expressive imprecise probability models than events do, as was shown by Walley [41, Section 2.7].[1] An event A is a subset of Ω, and it can be interpreted as a special gamble yielding one utile if it occurs and zero if it doesn't: an event A can be identified with its indicator I_A, which is a gamble.

In the behavioural approach to uncertainty modelling, a modeller's uncertainty about a possibility space Ω is measured through her attitudes to gambles defined on Ω. I mention two equivalent ways in which such attitudes can be represented mathematically: using coherent lower and upper previsions, and using sets of linear previsions.

2.1. **Lower and upper previsions.**

A modeller's uncertainty can be measured by eliciting her *lower prevision*, or supremum acceptable buying price, $\underline{P}(X)$ and her *upper prevision*, or infimum acceptable selling price, $\overline{P}(X)$ for gambles X. The transaction in which a gamble X is bought for a price x has reward function $X-x$, which is also a gamble. A modeller's *supremum acceptable buying price* $\underline{P}(X)$ for X is the largest real number c such that she is committed to accept the gamble $X-x$ for all $x<c$. At the same time, her *infimum acceptable selling price* $\overline{P}(X)$ for X is the smallest real number d such that she is committed to accept the gamble $x-X$ for all $x>d$. Clearly, buying a gamble X for price x has the same reward function as selling $-X$ for price $-x$, so it is a rationality requirement that $\overline{P}(-X)=-\underline{P}(X)$. This means that it is very easy to convert upper previsions into equivalent lower previsions, and vice versa. I shall do so regularly throughout the paper. In this section, I focus on lower previsions, but the conversion to upper previsions is straightforward to make.

[1]For *precise* probability models, it turns out to be immaterial which of the two approaches is chosen.

Recall that an *event* $A \subseteq \Omega$ can be identified with its indicator, and it turns out that buying and selling prices for its indicator I_A can be regarded as betting rates on and against A (I shall explain this in more detail in the next section). The lower prevision $\underline{P}(I_A)$ of the gamble I_A is also called the *lower probability* of A, and denoted simply as $\underline{P}(A)$. In a similar vein, upper probabilities are just upper previsions of (indicators of) events.[2]

Lower previsions represent a modeller's commitments to buy certain gambles. Buying gambles is a quite general model for behaviour in the face of uncertainty. Since we want such behaviour to be rational, a modeller's lower previsions should satisfy a number of consistency criteria. These can be summarised by the following requirement. Assume that the modeller specifies a lower prevision $\underline{P}(X)$ for all gambles X in some subset $\mathscr{K}$ of $\mathscr{L}(\Omega)$. In order to identify its possibility space and domain, I also denote this lower prevision by $(\Omega, \mathscr{K}, \underline{P})$.

Definition 1. The lower prevision $(\Omega, \mathscr{K}, \underline{P})$ is called *coherent* if for any natural numbers $n \geq 0$ and $m \geq 0$, and for any gambles $X_o, X_1, \ldots, X_n$ in $\mathscr{K}$:

$$\sup_{\omega\in\Omega}\left[\sum_{k=1}^{n}[X_k(\omega)-\underline{P}(X_k)] - m[X_o(\omega)-\underline{P}(X_o)]\right] \geq 0. \tag{1}$$

If the condition (1) holds in particular for the fixed choice $m = 0$, then $\underline{P}$ is said to *avoid sure loss*. The set of all coherent lower previsions on $\mathscr{L}(\Omega)$ is denoted by $\underline{\mathbb{P}}$.

If $\underline{P}$ avoids sure loss, there is no finite number of buying transactions, which the modeller is committed to accept as a result of her specifying $\underline{P}$, such that the net result of these combined transactions is always smaller than some strictly negative amount of utility. If $\underline{P}$ is coherent, the modeller cannot, by combining buying transactions that she finds acceptable, be induced to pay more for a gamble than she has specified in her lower prevision for it. See [41, Chapter 2] for a more detailed discussion and justification of these rationality criteria.

If $\mathscr{K}$ is a linear space of gambles, a coherent lower prevision can be characterised much more easily as follows [41, Theorem 2.5.5].

Theorem 2. *A lower prevision $\underline{P}$ defined on a linear subspace $\mathscr{K}$ of $\mathscr{L}(\Omega)$ is coherent if and only if for all X, Y in $\mathscr{K}$ and $\lambda > 0$:*

(i) $\underline{P}(X) \geq \inf X$ *[accepting sure gains];*
(ii) $\underline{P}(\lambda X) = \lambda \underline{P}(X)$ *[positive homogeneity];*
(iii) $\underline{P}(X+Y) \geq \underline{P}(X) + \underline{P}(Y)$ *[super-linearity]*

Throughout, I use the notation $\inf X$ for the infimum $\inf_{\omega\in\Omega} X(\omega)$ of a gamble X on its domain Ω, and similarly for $\sup X$.

So far, I have discussed lower previsions defined on subsets $\mathscr{K}$ of $\mathscr{L}(\Omega)$. The very important procedure of *natural extension* allows us to 'extend' a lower prevision $\underline{P}$ on $\mathscr{K}$ that avoids sure loss to a coherent lower prevision on all gambles by taking only two things into account: (a) the information contained in $\underline{P}$, and (b) the requirement of coherence. To see how, consider any gamble X on Ω. Assume that p is our modeller's supremum acceptable buying price for X. If this assessment is to be coherent with the lower prevision assessments $\underline{P}$, it is necessary that $p \geq \underline{E}(X)$, where

$$\underline{E}(X) = \sup_{n,\lambda_k,X_k} \inf_{\omega\in\Omega}\left[X(\omega) - \sum_{k=1}^{n}\lambda_k[X_k(\omega)-\underline{P}(X_k)]\right], \tag{2}$$

and the supremum runs over all integer $n \geq 0$, all real $\lambda_k \geq 0$ and all gambles $X_k \in \mathscr{K}$, for $k = 1, \ldots, n$. The lower prevision $\underline{E}$ defined by Eq. (2) is called the *natural extension* of

[2] In standard probability theory, the letter P is used for 'probability' (defined on events), and the letter E for 'expectation' (defined on gambles, or random variables). Here, I do not follow this practise: I use the symbol P to refer to 'prevision' (defined on *both* events and gambles), and 'E' further on to refer to 'extension', following [20] and [41], respectively.

the lower prevision $\underline{P}$. It is defined for any gamble X on Ω. Natural extension derives its importance from the following result, proved by Walley [41, Theorem 3.1.2].

Theorem 3 (Natural extension theorem)**.** *Let $\underline{P}$ be a lower prevision on a set of gambles $\mathscr{K} \subseteq \mathscr{L}(\Omega)$ that avoids sure loss, and let $(\Omega, \mathscr{L}(\Omega), \underline{E})$ be its natural extension. The following statements hold.*

- (i) $\inf X \leq \underline{E}(X)$ *for all* $X \in \mathscr{L}(\Omega)$.
- (ii) $\underline{E}$ *is a coherent lower prevision on* $\mathscr{L}(\Omega)$.
- (iii) $\underline{E}$ *dominates* $\underline{P}$ *on* $\mathscr{K}$ *:* $\underline{E}(X) \geq \underline{P}(X)$ *for all* $X \in \mathscr{K}$.
- (iv) $\underline{E}$ *coincides with* $\underline{P}$ *on* $\mathscr{K}$ *if and only if* $\underline{P}$ *is coherent.*
- (v) $\underline{E}$ *is the (point-wise) smallest coherent lower prevision on* $\mathscr{L}(\Omega)$ *that dominates* $\underline{P}$ *on* $\mathscr{K}$.
- (vi) *If* $\underline{P}$ *is coherent then* $\underline{E}$ *is the (point-wise) smallest coherent lower prevision on* $\mathscr{L}(\Omega)$ *that coincides with* $\underline{P}$ *on* $\mathscr{K}$.

This indicates that natural extension is *least-committal*: any coherent extension of the coherent lower prevision $\underline{P}$ implies a commitment to buy gambles X for a price that is at least as high as $\underline{E}(X)$, and it therefore has behavioural implications that are at least as strong. Moreover, if $\underline{P}$ is a lower prevision that avoids sure loss but is not coherent, natural extension corrects and extends it to a coherent lower prevision on all gambles, again with minimal behavioural implications.

Well-known examples of coherent lower probabilities are: (finitely additive) probability measures, 2-monotone set functions, completely monotone set functions (belief functions), and necessity measures. Their respective conjugate coherent upper probabilities are: (finitely additive) probability measures, 2-alternating set functions, completely alternating set functions (plausibility functions), and possibility measures. The natural extension of a probability measure to gambles turns out to be its Lebesgue integral, that of 2-monotone and 2-alternating set functions their Choquet integral.

2.2. **Linear previsions.** A modeller's supremum acceptable buying price and her infimum acceptable selling price for a gamble may differ because she is indecisive or because she has little information about the gamble. The difference between these buying and selling prices typically decreases as the amount of relevant information increases. In the special case that every gamble X has a 'fair price', meaning that the supremum acceptable buying price agrees with the infimum acceptable selling price, we obtain de Finetti's theory of (linear) previsions [20].

Definition 4. A *(linear) prevision P* on a set of gambles $\mathscr{K}$ is a map taking $\mathscr{K}$ to the set of real numbers $\mathbb{R}$, such that for all $m \geq 0$ and $n \geq 0$, and for any $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_m$ in $\mathscr{K}$,

$$\sup_{\omega \in \Omega} \left[\sum_{k=1}^{n} [X_k(\omega) - P(X_k)] - \sum_{k=1}^{m} [Y_k(\omega) - P(Y_k)] \right] \geq 0.$$

A linear prevision $(\Omega, \mathscr{K}, P)$ is therefore coherent, both when interpreted as a lower and as an upper prevision on $\mathscr{K}$. Linear previsions are the *precise* probability models, and they provide a link with the more traditional approaches to probability theory: a linear prevision defined on a field of events is simply a finitely additive probability measure. As already stated before, its natural extension to (measurable) gambles is nothing but the expectation associated with that measure, through Lebesgue (or Dunford [3]) integration.

If $\mathscr{K}$ is a linear space of gambles, a linear prevision can be characterised much more easily as follows.

Theorem 5. *A map P defined on a linear subspace $\mathscr{K}$ of $\mathscr{L}(\Omega)$ is a linear prevision if and only if for all X, Y in $\mathscr{K}$ and $\lambda \in \mathbb{R}$:*

- (i) $P(X) \geq \inf X$ *[accepting sure gains];*

(ii) $P(\lambda X) = \lambda P(X)$ *[homogeneity];*[3]
(iii) $P(X+Y) = P(X) + P(Y)$ *[linearity].*

In the rest of this paper, we are mainly interested in linear previsions that are defined on the set of all gambles $\mathscr{L}(\Omega)$. These can be characterised alternatively as real-valued maps P on $\mathscr{L}(\Omega)$ that (i) are linear: $P(\lambda X + \mu Y) = \lambda P(X) + \mu P(Y)$ for all gambles X and Y and real numbers λ and μ; (ii) are positive: $X \geq 0 \Rightarrow P(X) \geq 0$ for all gambles X; and (iii) have unit norm: $P(1) = 1$. The set of all linear previsions on $\mathscr{L}(\Omega)$ will be denoted by $\mathbb{P}$.

2.3. **Sets of linear previsions.** Linear previsions can be used to construct an alternative uncertainty model that is mathematically equivalent to a lower prevision: there is a very close relationship between a coherent lower prevision and a *set of* linear previsions. Indeed, consider a lower prevision $\underline{P}$ on the set of gambles $\mathscr{K}$. Define its *set of dominating linear previsions* by

$$\mathscr{M}(\underline{P}) = \{P \in \mathbb{P} \colon (\forall X \in \mathscr{K})(P(X) \geq \underline{P}(X))\}.$$

Then the following important result allows us to connect the models $\underline{P}$ and $\mathscr{M}(\underline{P})$ [41, Theorems 3.3.3 and 3.4.1].

Theorem 6 (Lower envelope theorem)**.** *Let $(\Omega, \mathscr{K}, \underline{P})$ be a lower prevision.*
(i) *$\underline{P}$ avoids sure loss if and only if $\mathscr{M}(\underline{P}) \neq \emptyset$.*
(ii) *$\underline{P}$ is coherent if and only if $\underline{P}(X) = \min\{P(X) \colon P \in \mathscr{M}(\underline{P})\}$ for all $X \in \mathscr{K}$, i.e., $\underline{P}$ is the* lower envelope *of $\mathscr{M}(\underline{P})$.*
(iii) *If $\underline{P}$ avoids sure loss, then the natural extension $\underline{E}$ of $\underline{P}$ is given by*

$$\underline{E}(X) = \min\{P(X) \colon P \in \mathscr{M}(\underline{P})\}$$

for all gambles X on Ω.

On the other hand, the lower envelope of any set of linear previsions is always a coherent lower prevision. This correspondence also establishes a link between the theory of coherent lower previsions and Bayesian sensitivity analysis [1].

Sometimes, both a lower prevision $\underline{P}$ and an upper prevision $\overline{P}$ are defined on a set of gambles $\mathscr{K}$. If these are *conjugate* in the sense that

$$\overline{P}(X) = -\underline{P}(-X), \quad X \in \mathscr{K} \cap (-\mathscr{K}),$$

where $-\mathscr{K} = \{-X \colon X \in \mathscr{K}\}$, then we can replace the pair $(\underline{P}, \overline{P})$ by a single lower prevision $\underline{P}'$ defined on $\mathscr{K} \cup (-\mathscr{K})$ by $\underline{P}'(X) = \underline{P}(X)$ if $X \in \mathscr{K}$ and $\underline{P}'(X) = -\overline{P}(-X)$ if $X \in -\mathscr{K}$. We then say that the pair $(\underline{P}, \overline{P})$ avoids sure loss (is coherent) if $\underline{P}'$ does (is). The natural extension of the pair $(\underline{P}, \overline{P})$ is simply the natural extension of $\underline{P}'$; and we write

$$\mathscr{M}(\underline{P}, \overline{P}) = \mathscr{M}(\underline{P}') = \{P \in \mathbb{P} \colon (\forall X \in \mathscr{K})(\underline{P}(X) \leq P(X) \leq \overline{P}(X))\}.$$

2.4. **Topological considerations.** The set of all gambles $\mathscr{L}(\Omega)$, provided with the topology of uniform convergence (induced by the supremum norm $\|X\| = \sup|X|$), is a Banach space (a complete, normed linear space). Its topological dual $\mathscr{L}(\Omega)^*$ is the collection of all linear functionals on $\mathscr{L}(\Omega)$, i.e., real-valued linear maps on $\mathscr{L}(\Omega)$, that are continuous with respect to this topology (and the usual topology on the reals). All linear previsions are continuous linear functionals on $\mathscr{L}(\Omega)$, and therefore $\mathbb{P} \subseteq \mathscr{L}(\Omega)^*$.

The *weak* topology* on $\mathscr{L}(\Omega)^*$ is the weakest topology that makes all so-called *evaluation functionals* $X^* \colon \mathscr{L}(\Omega)^* \to \mathbb{R}$ continuous, where $X \in \mathscr{L}(\Omega)$ and $X^*(\Lambda) = \Lambda(X)$ for all $\Lambda \in \mathscr{L}(\Omega)^*$. It is the topology of point-wise convergence on $\mathscr{L}(\Omega)^*$. It follows from Tychonov's theorem that $\mathbb{P}$ is weak*-compact, i.e., compact in this weak* topology. For any lower prevision $\underline{P}$ the set of dominating linear previsions $\mathscr{M}(\underline{P})$ is a weak*-closed subset of the weak*-compact $\mathbb{P}$, and therefore weak*-compact as well. It is clearly also convex.

[3]Actually, the second condition is redundant.

Given a gamble X in $\mathscr{L}(\Omega)$, we denote by $\mathscr{N}(X)$ the weak*-closed (and therefore weak*-compact) set $\{P \in \mathbb{P} \colon P(X) = 0\}$ of all linear previsions that map X to zero. It is the intersection of $\mathbb{P}$ with the kernel of the evaluation functional X^*.

3. JUSTIFYING POSSIBILISTIC PREVISIONS

Consider an unstable radioactive nucleus. Its probability of decay in a given time interval t is given by $1-e^{-\lambda t}$, where the parameter λ is the *decay rate* of the nucleus, a constant which could in principle be determined from the physical properties of the nucleus, such as its composition, excitation, … In practise, however, decay rates are determined experimentally, within a certain margin for error.

This is a special case of a more general kind of situation that occurs very often: an event A (or a random variable), has a *true* probability $P_T(A)$ (or a probability measure P_T), but there is not enough information to determine it unequivocally. An often used solution to this problem is the specification of a class of probability measures, of which the true unknown probability is a member. Evidently, the "smaller" the class, the better we know the probability. In the above example, this would amount to specifying an interval of values which we expect the true decay rate to belong to. A number of techniques have been developed for dealing with this kind of situation; see [1, 2, 34, 41] for more details and further references.

In many cases, however, it seems difficult to draw the line between the probability measures or distributions that will be included in the class, and the ones that have to be excluded from consideration. In the case described above, for instance, the precise choice of the lower and upper bounds for the interval of possible decay rates seems to some extent arbitrary. At the same time, not all members of the class will deserve the same status, as it may be felt that central members are more to be trusted than the ones near the border.

Another example should make this more clear. It is an adaptation of the "Miss Julie takes a bet" example, due to Gärdenfors and Sahlin [26]. Imagine a coin with a chance (or true probability) θ of landing 'tails'. A friend gives the modeller the following information: he has loaded the coin so heavily that it will almost always fall with the same side upward, only, he will not tell the modeller which side. So the modeller knows that θ is close to either 0 or 1, but she doesn't know which. She could model this by specifying a set of possible values for θ of the form

$$A_\alpha = \big[0, \frac{1-\alpha}{2}\big) \cup \big(\frac{1+\alpha}{2}, 1\big],$$

where α is some specific element of $[0,1]$. The fact that these sets are symmetric with respect to the possible value $\frac{1}{2}$ reflects the symmetry in the available information.

To give another example, her friend might have told the modeller that the chance θ for the coin to land 'tails' is close to 1, and she could model this by specifying a set of possible values for θ of the form

$$B_\alpha = (\alpha, 1],$$

where, again, α is some specific element of $[0,1]$.

In either case, the information the modeller has does not allow her to really choose between the different possible values of α. To a certain extent, this choice will be arbitrary. On the other hand, the smaller α, the greater A_α (or B_α), so the more she can be confident that the real θ will belong to A_α (or B_α).

The use of second-order or even higher-order probabilities [30, 31] has been suggested as a solution to this problem. Rather than a class of candidate probability measures, a probability measure is given on the class of all probability measures relevant to the problem at hand. In the coin tossing problem described above, this would amount to specifying a probability measure on the set $[0,1]$ of all possible values for θ.

There are a number of problems with this approach, however. First of all, the class of relevant probability measures may be very large, and there may be mathematical problems associated with the specification of reasonable probability measures on such classes. At the same time, it seems a bit strange that the available information would allow the modeller to specify a unique probability measure on a large and abstract higher-order space of probability measures, given that it does not allow her to pinpoint a single probability measure on the often much more tangible lower-order spaces. For a much more detailed discussion and critique, see [36, 41].

It seems more reasonable to use an imprecise probability model to describe the uncertainty about the true probability. Let me give an outline of how this can be done in the examples introduced above.

Let us return to our coin tossing problem, where the available information is that the chance θ is close to either 0 or 1. As we have seen, the available information leads the modeller to consider the events A_α, $\alpha \in [0,1]$. It provides some evidence for the chance θ belonging to these sets, but no evidence at all that θ belongs to their complements. In terms of behaviour, the given information seems to warrant her accepting bets *on* the events A_α, or *against* the events $\mathrm{co}A_\alpha$, and at higher rates the smaller α is. But it does *not* lead her to accept bets against the A_α, or on the $\mathrm{co}A_\alpha$.

Let me make this more clear, keeping in mind the ideas and notions behind lower previsions and probabilities introduced in the previous section. Consider the following type of gamble I_A involving an event A. If A occurs, the modeller gains one unit of utility, and if it doesn't, she gains nothing. If she is willing to pay a price r (in units of utility) to participate in this bet, this means that she is willing to engage in a gamble which yields $1-r$ if A occurs, and $-r$ if it doesn't. The number r is called a *betting rate*. Evidently, the more information the modeller has which leads her to believe that A will occur, the higher (closer to 1) will be the rate r at which she is willing to bet on A. If we observe that her *lower probability* of A is nothing but her supremum rate for betting on A, i.e., the number $\underline{P}(A)$ such that she is willing to bet on A at any rate $r < \underline{P}(A)$, this tells us that the more her evidence points in the direction of the occurrence of A, the higher her lower probability of A will be. If she has little or no evidence in favour of A, her lower probability $\underline{P}(A)$ will be close to zero.

On the other hand, evidence *against* the occurrence of A, or equivalently, in favour of the opposite event $\mathrm{co}A$, tends to increase her lower probability $\underline{P}(\mathrm{co}A)$ of $\mathrm{co}A$, or decrease her *upper probability* $\overline{P}(A) = 1 - \underline{P}(\mathrm{co}A)$ of A. Again, if she has little or no evidence against A, the modeller's upper probability $\overline{P}(A)$ will be close to one. Ignorance, or little or no evidence either way, tends to produce upper and lower probabilities with high *imprecision* $\overline{P}(A) - \underline{P}(A)$. More evidence tends to drive $\underline{P}(A)$ and $\overline{P}(A)$ closer together. As we also indicated in the previous section, if $\underline{P}(A) = \overline{P}(A)$, this common value is called the *probability* $P(A)$ of A, and it can be interpreted as a *fair betting rate* in de Finetti's sense [20, 21]: if $r < P(A)$ the modeller accepts bets on A at rate r and if $r > P(A)$ she accepts bets against A at rate $1-r$.

The main idea behind all this is that evidence in favour of an event is modelled by its lower probability, whereas evidence against it is modelled by its upper probability. This, by the way, is also a central idea in Shafer's theory of evidence [40]. In our example, it therefore seems natural to model the available information by an upper probability assessment $\overline{P}$ on the sets $\mathrm{co}A_\alpha$ for $\alpha \in [0,1]$—or equivalently, a lower probability assessment on the A_α—in the following way:

$$\overline{P}(\mathrm{co}A_\alpha) = g(\alpha), \tag{3}$$

where g is some continuous, non-decreasing transformation of $[0,1]$ with $g(1) = 1$. It can be shown, using the results in [15], that such an assessment is—through the procedure of natural extension explained in the previous section—behaviourally equivalent to the following upper probability $\overline{P}$ on the class of all subsets B of $[0,1]$:

$$\overline{P}(B) = \sup_{\vartheta \in B} g(|1 - 2\vartheta|),$$

This $\overline{P}$ is special in that it takes any union $\bigcup_{j\in J} A_j$ of events into a supremum:

$$\overline{P}(\bigcup_{j\in J} A_j) = \sup_{j\in J} \overline{P}(A_j).$$

An upper probability with this property is called a *possibility measure* in the literature [5, 6, 7, 24, 25, 50]. It is completely characterised by its so-called *distribution* $\pi\colon [0,1] \to [0,1]$, which summarises the values $\overline{P}$ takes on singletons:

$$\pi(\vartheta) = \overline{P}(\{\vartheta\}) = g(|1-2\vartheta|), \quad \vartheta \in [0,1],$$

and $\overline{P}(B) = \sup_{\vartheta\in B} \pi(\vartheta)$. Observe that the possibility measure $\overline{P}$ and its distribution π are *normal*, in that $\overline{P}([0,1]) = \sup_{\vartheta\in[0,1]} \pi(\vartheta) = 1$. The upper probability $\overline{P}$ represents the minimal behavioural implications of the given assessments (3), and therefore seems a good model for the given information.

Similarly, the information that θ is close to 1 leads us to consider the events B_α, $\alpha \in [0,1]$, and it seems to warrant accepting bets against the events $\mathrm{co}B_\alpha$, and at higher rates the smaller α is. Again, it seems natural to model the available information by upper probability assessments $\overline{P}$ on the events $\mathrm{co}B_\alpha$ for $\alpha \in [0,1]$ in the following way:

$$\overline{P}(\mathrm{co}B_\alpha) = f(\alpha), \tag{4}$$

where f is some continuous, non-decreasing transformation of $[0,1]$ with $f(1) = 1$. This assessment is again behaviourally equivalent—in the sense of natural extension—to an upper probability $\overline{P}$ that is a normal possibility measure with distribution

$$\pi(\vartheta) = f(\vartheta), \quad \vartheta \in [0,1].$$

In other words, for any $B \subseteq [0,1]$:

$$\overline{P}(B) = \sup_{\vartheta\in B} f(\vartheta).$$

What makes these examples interesting is that they—and the reasoning behind them—are quite typical. *The crucial point is that the available linguistic information allows a modeller to specify in a natural way an upper (or lower) probability assessment on a class of nested sets, and this will fairly often be the case, especially if the given information is vague*. A more detailed discussion of these issues can be found in [43]. Given a number of additional continuity assumptions, it can be shown that the natural extension is a possibility measure [15, 43].

For this reason, imprecise second-order probability models, where the second-order upper probability is a normal possibility measure, seem rather important. Interestingly, a thorough study of precisely this type of second-order uncertainty models is the subject of a fairly recent paper by Walley [42]. It describes how the second-order imprecise model can, through natural extension, be converted into a first-order imprecise probability model, which can then be used as prior information for decision making and statistical reasoning, according to the general techniques described in [41].

In the present paper, I discuss a number of interesting aspects of this model that are left untouched in the above-mentioned paper, and that provide additional justifications for working with second-order possibility measures (or distributions).

4. REPRESENTABILITY: FROM A LOCAL TO A GLOBAL MODEL

Consider a situation where there is uncertainty about some phenomenon of interest, or to use a different terminology, about the state of the world ω in Ω. I shall first assume that this uncertainty is described by some 'ideal' probability measure—or in the language of Section 2—by some linear prevision P_T.

I am well aware that assuming the existence of an ideal or true precise prevision P_T is questionable. But I shall show in Section 8 that this assumption is not necessary: the material in Sections 5–7 can be developed without it, using the behavioural notion of a price function. The reason why I prefer to first formulate my model in a manner that is based on this assumption, is purely didactic: it allows me to stay much closer to the existing Bayesian hierarchical models and fuzzy probability models in the literature. Starting straight away with price functions functions would leave the reader in the dark about the relevant and interesting connections with these other models. So let us, at least for the time being, ignore any qualms we might have about 'true probabilities', and proceed as if there really were such a thing as a true or ideal prevision P_T. It could be a physical or objective probability, as in the case of a decay rate, or it could be somebody's subjective Bayesian model that the modeller does not know exactly: for the coin tossing example in the previous section, it could be our friend's personal probability for tails that he is giving the modeller linguistic information about ...

Consider a gamble X on the possibility space Ω, and suppose that a modeller has information about its true prevision (or expectation) $P_T(X)$ in the form of a linguistic assessment, such as "the prevision for X is 10", or "... close to 5", or "... between 2 and 8", or "... fairly high". I have argued in the previous section that this information can in many cases be modelled adequately by a normal possibility distribution on the set of possible values for $P_T(X)$—the set of real numbers, or in the case that X is (the indicator of) an event A, the unit interval $[0,1]$. I shall call this possibility distribution the modeller's *possibilistic prevision for X* and denote it by $\mathfrak{p}(X)$. For an event A, the possibilistic prevision $\mathfrak{p}(I_A)$ is also denoted by $\mathfrak{p}(A)$ and called the *possibilistic probability* of A. As mentioned before, I want to consider the more general model of gambles and previsions rather than just events and probabilities, because it is so much more expressive.

Remark 1. In the present work, as well as in previous papers [8, 28], I use the name *possibilistic* rather than *fuzzy* prevision, because I think it is the better name: it is a *possibility* distribution representing information about the value of an ideal *prevision*. Although formally speaking, any normal possibility distribution is a fuzzy set in the sense of Zadeh [46], it seems to me that there is a very important difference between the two as far as their interpretation is concerned: it was shown in [15, 16, 43] that a normal possibility distribution has a very clear behavioural interpretation in terms of upper betting rates, whereas I have yet to see a convincing behavioural interpretation for a fuzzy set (or its so-called membership function). See [43] for a more detailed exposition of my point of view. We shall see later in this section (Remark 2) and in the next section (Remark 3) that fuzzy probabilities have a different, and from my point of view less defensible, calculus as well. $\Diamond$

There is nothing that prevents the modeller from considering vague or clear linguistic assessments of the true prevision for more than one gamble. Suppose that she has such assessments for all gambles X in a subset $\mathscr{K}$ of $\mathscr{L}(\Omega)$, and that for each X in $\mathscr{K}$, she models the corresponding assessment by a normal possibility distribution $\mathfrak{p}(X)$ on the set of real numbers, i.e., $\mathfrak{p}(X)$ is a map from $\mathbb{R}$ to $[0,1]$ such that $\sup_{x\in\mathbb{R}}\mathfrak{p}(X)\cdot x=1$, where we use the notation $\mathfrak{p}(X)\cdot x$ to denote the value of the map $\mathfrak{p}(X)$ in the real number x. The modeller has thus represented the available information by a map $\mathfrak{p}$, defined on the set $\mathscr{K}$, whose values $\mathfrak{p}(X)$, $X\in\mathscr{K}$ are normal possibility distributions on the reals. I shall call such a map a *possibilistic prevision*.

Definition 7. A *possibilistic prevision* $\mathfrak{p}$ on an arbitrary subset $\mathscr{K}$ of $\mathscr{L}(\Omega)$ is a map taking the gambles X in $\mathscr{K}$ to normal possibility distributions $\mathfrak{p}(X)$ on $\mathbb{R}$. In order to clearly identify its possibility space and domain, I shall often denote such a possibilistic prevision by $(\Omega,\mathscr{K},\mathfrak{p})$.

In order to avoid all possible confusion at this point, I want to stress again that $\mathfrak{p}(X)$ is a possibility distribution on $\mathbb{R}$. Its value $\mathfrak{p}(X)\cdot x$ in a point $x\in\mathbb{R}$ can be interpreted as the modeller's upper betting rate for the event that the true prevision $P_T(X)$ for the gamble X is equal to x. A few examples will clarify this.

Example 1. How can we represent the fact that the modeller has no information at all about the value $P_T(X)$ of the true prevision P_T in a gamble X? It follows from the properties of a linear prevision (see Lemma 42 in Appendix A) that $P_T(X)$ must be an element of the real interval $[\inf X, \sup X]$, and since the modeller knows nothing else about $P_T(X)$, it can in principle assume any value in this interval. If A is a subset of the reals, this means that the modeller should not be prepared to bet against A at any (non-trivial) odds if $A \cap [\inf X, \sup X] \neq \emptyset$, because for all she knows, $P_T(X)$ may assume any value in $[\inf X, \sup X]$, and might therefore well lie inside A. On the other hand, if $[\inf X, \sup X] \subseteq \mathrm{co}A$, she should be prepared to bet against A at all odds, because she knows for sure that $P_T(X)$ belongs to $\mathrm{co}A$. She may therefore model the available information by the *vacuous upper probability* $\overline{P}$ *relative to* $[\inf X, \sup X]$, given by, for any $A \subseteq \mathbb{R}$:

$$\overline{P}(A) = \begin{cases} 1 & \text{if } A \cap [\inf X, \sup X] \neq \emptyset \\ 0 & \text{if } A \cap [\inf X, \sup X] = \emptyset. \end{cases}$$

This is a possibility measure on $\mathbb{R}$ whose distribution will be denoted by $\mathfrak{p}_v(X)$: for any $x \in \mathbb{R}$,

$$\mathfrak{p}_v(X) \cdot x = \begin{cases} 1 & \text{if } x \in [\inf X, \sup X] \\ 0 & \text{elsewhere}, \end{cases} \tag{5}$$

and for any $A \subseteq \mathbb{R}$, $\overline{P}(A) = \sup_{x \in A} \mathfrak{p}_v(X) \cdot x$. Eq. (5) defines a possibilistic prevision, called the *vacuous possibilistic prevision* on $\mathscr{L}(\Omega)$. It represents the fact that the modeller has no information whatsoever about the true model P_T. Its restriction to a set of gambles $\mathscr{K}$ will be called the vacuous possibilistic prevision on $\mathscr{K}$. ♦

Example 2. Consider the assessment discussed in the previous section: 'the probability that this coin will land tails on the next toss is very close to 1'. How can this assessment be modelled?

If we assume for the sake of simplicity that the coin can only land heads or tails, we may write for the possibility space $\Omega = \{h, t\}$. I have argued extensively in the previous section that the given assessment is behaviourally equivalent to specifying a possibilistic prevision (probability) $\mathfrak{p}$ on the set of events $\mathscr{K} = \{\{t\}\}$, where

$$\mathfrak{p}(\{t\}) \cdot x = \begin{cases} f(x) & \text{if } x \in [0,1] \\ 0 & \text{elsewhere}, \end{cases}$$

where f is defined through the assessments (4). ♦

Example 3. A similar course of reasoning can be followed for the assessment: 'the probability that this coin will land tails on the next toss is close to 1 or close to 0'. It was also argued in the previous section that this linguistic assessment can be modelled by a possibilistic prevision $\mathfrak{p}$ on $\mathscr{K} = \{\{t\}\}$, where

$$\mathfrak{p}(\{t\}) \cdot x = \begin{cases} g(|2x-1|) & \text{if } x \in [0,1] \\ 0 & \text{elsewhere} \end{cases}$$

where g is defined through the assessments (3). ♦

Recall that the modeller is assuming that the uncertainty about the phenomenon of interest is modelled by some true linear prevision P_T, which is in principle defined on all gambles X on Ω. So far, we have been concerned with representing linguistic information about the *local* aspects of the true model, i.e., about the *values* $P_T(X)$ that this true linear prevision P_T assumes in certain gambles X. In this sense, a possibilistic prevision $\mathfrak{p}$ could be called a *local model*.

But, rather than concentrating on its values in gambles, we could also pay attention to the map P_T itself. As first suggested by Walley in [42], we could for instance consider a normal (so-called second-order) possibility distribution π on the set $\mathbb{P}$ of all linear previsions, whose interpretation is that $\pi(P)$ is the modeller's upper probability for the event that the linear prevision $P \in \mathbb{P}$ is the true prevision P_T. More generally, the upper probability for the event that the true prevision belongs to a subset A of $\mathbb{P}$ is then given by $\sup\{\pi(P) \colon P \in A\}$.[4] Such a normal second-order possibility distribution π could for obvious reasons be called a *global* model.

Any global model π leads in a straightforward manner to a local one. Indeed, consider an arbitrary gamble X on Ω, and a real number x. Then $\mathfrak{p}(X) \cdot x$, the modeller's upper probability that the true prevision $P_T(X)$ for X is equal to x, is nothing but the upper probability of the set of linear previsions $\mathscr{N}(X-x) = \{P \in \mathbb{P} \colon P(X) = x\}$:

$$\mathfrak{p}(X) \cdot x = \sup\{\pi(P) \colon P \in \mathscr{N}(X-x)\} = \sup\{\pi(P) \colon P(X) = x\}. \tag{6}$$

The question that will occupy us in the rest of this section and in Section 5, is whether we can go in the opposite direction: can we derive a global model from a given local one? Before addressing it, however, let me discuss my reasons for wanting to do so.

It should be clear that a local model is much easier to obtain than a global one. It is arguably conceptually easier, and it requires much less effort, to concentrate on the true prevision of a (possibly small) number of gambles rather than to concentrate on P_T itself, or what is equivalent, to concentrate on its values in *every* conceivable gamble X on Ω. Moreover, the process of assessment naturally leads to a local model.

But, on the other hand, if we want to use the assessments as a starting point for decision-making, or as prior information in statistical reasoning, it seems that we need a global model. Walley has indeed shown [42] how we can use *coherence arguments* to derive a first-order uncertainty model from a normal second-order possibility distribution (a global model)—in essentially the same way as a precise second-order probability leads to a precise first-order probability—and how this first-order model can be used in decision making and statistical reasoning. I shall devote more attention to this in Section 6.

A third reason is that we shall see in this and the next section that connecting local with global models allows us to establish a very close formal link between possibilistic previsions and Walley's behavioural theory of imprecise probabilities [41].

So, to conclude, it seems that in order to be able to use the assessments embodied in the *local* model $\mathfrak{p}$ for the practical purposes of reasoning and decision-making, it will be a great step forward if we are able to convert it into a *global*, normal second-order possibility distribution. This leads to the criterion of representability, which is the consistency criterion that I shall want a possibilistic prevision to satisfy.

Definition 8. A possibilistic prevision $(\Omega, \mathscr{K}, \mathfrak{p})$ is called *representable* if there is a normal possibility distribution $\pi \colon \mathbb{P} \to [0,1]$ that *represents* $\mathfrak{p}$, i.e., such that for all $x \in \mathbb{R}$ and $X \in \mathscr{K}$:

$$\mathfrak{p}(X) \cdot x = \sup\{\pi(P) \colon P(X) = x\}. \tag{7}$$

Any π that represents $\mathfrak{p}$ will be called a *representation* of $\mathfrak{p}$. $(\Omega, \mathscr{K}, \mathfrak{p})$ is called *reasonable* if there is a normal possibility distribution $\pi \colon \mathbb{P} \to [0,1]$ such that for all $x \in \mathbb{R}$ and $X \in \mathscr{K}$:

$$\mathfrak{p}(X) \cdot x \geq \sup\{\pi(P) \colon P \in \mathbb{P} \text{ and } P(X) = x\}.$$

Observe that reasonability is necessary for representability: any representable possibilistic prevision is in particular reasonable.

If we consider the set function μ defined on the collection $\mathscr{A}$ of subsets $\mathscr{N}(X-x) = \{P \in \mathbb{P} \colon P(X) = x\}$ of $\mathbb{P}$ by $\mu(\mathscr{N}(X-x)) = \mathfrak{p}(X) \cdot x$, for all $x \in \mathbb{R}$ and $X \in \mathscr{K}$, then it is obvious that (i) $\mathfrak{p}$ is reasonable if and only if there is a normal possibility measure on $\mathbb{P}$ that is dominated by μ on its domain $\mathscr{A}$; and (ii) $\mathfrak{p}$ is representable if and only if the set function μ can be extended to a normal possibility measure on $\mathbb{P}$. The well-known Possibilistic Extension Theorem [4, 44] (see Theorem 41 in Appendix A) tells us that in

[4] I am concentrating here on global models that are possibility measures, as they are sufficiently general within the context of this paper. For a discussion of more general models, see [9, 13].

order to find out whether $\mathfrak{p}$ is reasonable/representable, we must look at the possibility distribution π^g on $\mathbb{P}$, given by

$$\pi^g(P) = \inf_{A \in \mathscr{A}, P \in A} \mu(A) = \inf_{\substack{X \in \mathscr{K}, x \in \mathbb{R} \\ P(X)=x}} \mathfrak{p}(X) \cdot x = \inf_{X \in \mathscr{K}} \mathfrak{p}(X) \cdot P(X).$$

This possibility distribution π^g on $\mathbb{P}$ is completely determined by the values of the possibilistic prevision $\mathfrak{p}$ on its domain $\mathscr{K}$, and it will play a central part in this paper. I shall therefore give it the special notation $\mathscr{M}(\mathfrak{p})$. The connection with the previously introduced notation $\mathscr{M}(\underline{P})$ will become clear further on.

Definition 9. With any possibilistic prevision $(\Omega, \mathscr{K}, \mathfrak{p})$ we can associate a possibility distribution $\mathscr{M}(\mathfrak{p})$ on $\mathbb{P}$, defined by:

$$\mathscr{M}(\mathfrak{p}) \colon \mathbb{P} \to [0,1] \colon P \mapsto \mathscr{M}(\mathfrak{p}) \cdot P = \inf_{X \in \mathscr{K}} \mathfrak{p}(X) \cdot P(X).$$

The following theorem is an immediate translation of the Possibilistic Extension Theorem to the problem under study here. It shows that the second-order possibility distribution $\mathscr{M}(\mathfrak{p})$ plays a central part in the representation issue: if the possibilistic prevision $\mathfrak{p}$ has a representation, then the greatest, and therefore least-committal, or most conservative, one is precisely $\mathscr{M}(\mathfrak{p})$.

Theorem 10 (Representation theorem)**.** *Let $(\Omega, \mathscr{K}, \mathfrak{p})$ be a possibilistic prevision. Then the following statements hold.*

(*i*) *$\mathscr{M}(\mathfrak{p})$ is the point-wise greatest possibility distribution π on $\mathbb{P}$ such that for all $X \in \mathscr{K}$ and $x \in \mathbb{R}$, $\sup\{\pi(P) \colon P(X) = x\} \leq \mathfrak{p}(X) \cdot x$.*

(*ii*) *$\mathfrak{p}$ is reasonable if and only if $\mathscr{M}(\mathfrak{p})$ is a normal possibility distribution on $\mathbb{P}$, i.e., if*

$$\sup_{P \in \mathbb{P}} \mathscr{M}(\mathfrak{p}) \cdot P = \sup_{P \in \mathbb{P}} \inf_{X \in \mathscr{K}} \mathfrak{p}(X) \cdot P(X) = 1.$$

(*iii*) *If $\mathfrak{p}$ is reasonable, then $\mathfrak{p}$ is representable if and only if $\mathfrak{p}$ has representation $\mathscr{M}(\mathfrak{p})$, i.e, if for all $X \in \mathscr{K}$ and $x \in \mathbb{R}$,*

$$\mathfrak{p}(X) \cdot x = \sup_{P(X)=x} \mathscr{M}(\mathfrak{p}) \cdot P = \sup_{P(X)=x} \inf_{Y \in \mathscr{K}} \mathfrak{p}(Y) \cdot P(Y).$$

(*iv*) *If $\mathfrak{p}$ is representable, then $\mathscr{M}(\mathfrak{p})$ is its greatest representation: for any representation π of $\mathfrak{p}$, $\pi(P) \leq \mathscr{M}(\mathfrak{p}) \cdot P$ for all $P \in \mathbb{P}$.*

Example 4 (Continuation of Example 1). The vacuous possibilistic prevision $\mathfrak{p}_v$ on a non-empty set of gambles $\mathscr{K}$ is always representable. Its greatest representation $\mathscr{M}(\mathfrak{p}_v)$ is given by

$$\mathscr{M}(\mathfrak{p}_v) \cdot P = \inf_{X \in \mathscr{K}} \mathfrak{p}_v(X) \cdot P(X) = 1$$

for all $P \in \mathbb{P}$, taking into account Lemma 42 in Appendix A. Note that this holds independently of the choice of $\mathscr{K}$. The possibility measure on $\mathbb{P}$ with distribution $\mathscr{M}(\mathfrak{p}_v)$ is the vacuous upper probability: it assumes the value 1 in any non-empty subset of $\mathbb{P}$, which reflects the fact that we have no information about which value P_T assumes in $\mathbb{P}$. ♦

We can use Theorem 10 to prove a number of interesting properties for representable possibilistic previsions.

Proposition 11. *Let $(\Omega, \mathscr{K}, \mathfrak{p})$ be a representable possibilistic prevision. The following properties hold whenever the gambles involved are in $\mathscr{K}$, and μ, λ and x are real numbers.*[5]

(*i*) *$\mathfrak{p}(X) \cdot x = 0$ if $x < \inf X$ or $x > \sup X$, or equivalently, $\mathfrak{p}(X) \leq \mathfrak{p}_v(X)$.*

(*ii*) *$\mathfrak{p}(\mu) \cdot x = 1$ if $x = \mu$, and $\mathfrak{p}(\mu) \cdot x = 0$ if $x \neq \mu$.*

---

[5]These properties imply that for an event A, $\mathfrak{p}(A)$ is zero outside the unit interval $[0,1]$, and that for all real x, $\mathfrak{p}(\mathrm{co}A) \cdot x = \mathfrak{p}(A) \cdot (1-x)$.

(*iii*) *if* $\lambda \neq 0$ *then* $\mathfrak{p}(\lambda X)\cdot x = \mathfrak{p}(X)\cdot(x/\lambda)$.
(*iv*) $\mathfrak{p}(-X)\cdot x = \mathfrak{p}(X)\cdot(-x)$.
(*v*) $\mathfrak{p}(X)\cdot x = \mathfrak{p}(\mu+X)\cdot(\mu+x) = \mathfrak{p}(\mu-X)\cdot(\mu-x)$.
(*vi*) $\mathfrak{p}(X+Y)\cdot x \leq \sup_{y\in\mathbb{R}}\min\{\mathfrak{p}(X)\cdot y, \mathfrak{p}(Y)\cdot(x-y)\}$.

Proof. Recall that it follows from the representability of $\mathfrak{p}$ that for any gamble X in the domain of $\mathfrak{p}$, $\mathfrak{p}(X)\cdot x = \sup\{\mathscr{M}(\mathfrak{p})\cdot P\colon P\in\mathscr{N}(X-x)\}$, and that $\mathscr{M}(\mathfrak{p})$ is a normal $\mathbb{P}-[0,1]$-map. To prove the first property, note that by Lemma 42, $\mathscr{N}(X-x)=\emptyset$ for x outside $[\inf X, \sup X]$, whence $\mathfrak{p}(X)\cdot x = 0$. To verify the second property, note that for any $x\neq\mu$, $\mathscr{N}(\mu-x)=\emptyset$, so $\mathfrak{p}(\mu)\cdot x = 0$, whereas $\mathscr{N}(\mu-\mu)=\mathbb{P}$, whence $\mathfrak{p}(\mu)\cdot\mu = 1$, since $\mathscr{M}(\mathfrak{p})$ is normal. To prove the third property, it suffices to point out that if $\lambda\neq 0$, then $\mathscr{N}(\lambda X - x) = \mathscr{N}(X-(x/\lambda))$, so that $\mathfrak{p}(\lambda X)\cdot x = \mathfrak{p}(X)\cdot(x/\lambda)$. Obviously the third property implies the fourth, by letting $\lambda=-1$. The fifth property follows from $\mathscr{N}(X-x) = \mathscr{N}((\mu+X)-(\mu+x)) = \mathscr{N}((\mu-X)-(\mu-x))$, whence $\mathfrak{p}(X)\cdot x = \mathfrak{p}(\mu+X)\cdot(\mu+x) = \mathfrak{p}(\mu-X)\cdot(\mu-x)$. To prove the last property, observe that $\mathscr{N}(X+Y-x) = \bigcup_{y\in\mathbb{R}}\mathscr{N}(X-y)\cap\mathscr{N}(Y-x+y)$, whence

$$\mathfrak{p}(X+Y)\cdot x = \sup_{y\in\mathbb{R}}\sup_{P\in\mathscr{N}(X-y)\cap\mathscr{N}(Y-x+y)}\mathscr{M}(\mathfrak{p})\cdot P.$$

Since

$$\sup_{P\in\mathscr{N}(X-y)\cap\mathscr{N}(Y-x+y)}\mathscr{M}(\mathfrak{p})\cdot P \leq \sup_{P\in\mathscr{N}(X-y)}\mathscr{M}(\mathfrak{p})\cdot P = \mathfrak{p}(X)\cdot y$$

and similarly

$$\sup_{P\in\mathscr{N}(X-y)\cap\mathscr{N}(Y-x+y)}\mathscr{M}(\mathfrak{p})\cdot P \leq \sup_{P\in\mathscr{N}(Y-x+y)}\mathscr{M}(\mathfrak{p})\cdot P = \mathfrak{p}(Y)\cdot(x-y),$$

this completes the proof. □

Remark 2. These properties, and in particular the addition rule (vi), show that representable possibilistic previsions have different properties than the fuzzy probabilities introduced by Zadeh (see [52] for a recent discussion), although they are formally related, and try to model the same thing: vague probability assessments. I discuss the differences between the two models in more detail in the next section, see Remark 3.

My reasons for preferring my model are twofold. First of all, representability is *useful*. As we shall see in Section 6, it is the representable possibilistic previsions that we shall be able to use in decision making and statistical reasoning, using arguments that are based on coherence, i.e., rationality requirements. Secondly, there is a definite sense in which representable possibilistic previsions are also *necessary*. Indeed, representability is a direct consequence of the following two assumptions: (i) that there is a true (but ill-known to the modeller) prevision P_T, and (ii) that the modeller's information can be represented by a possibility distribution on the set of all linear previsions. Then the corresponding local model *should* be, i.e., can be nothing else but a representable possibilistic prevision. Anyone wanting to use other models, such as possibilistic previsions with different properties—and in particular fuzzy probabilities with different addition rules—would have to let go of at least one of the two above-mentioned assumptions. ◊

Representability is clearly a desirable property for a possibilistic prevision to have. It turns out that if a possibilistic prevision $(\Omega,\mathscr{K},\mathfrak{p})$ is not representable but only reasonable, there is an interesting way to correct it into a representable possibilistic prevision. Indeed, for a reasonable $\mathfrak{p}$, the associated second-order possibility distribution $\mathscr{M}(\mathfrak{p})$ is normal, so we can use it to generate a new possibilistic prevision defined on all gambles, using the procedure summarised in Eq. (6).

Definition 12. Let $(\Omega,\mathscr{K},\mathfrak{p})$ be a reasonable possibilistic prevision, and let $\mathscr{M}(\mathfrak{p})$ be its associated normal second-order possibility distribution. Then the possibilistic prevision

$(\Omega, \mathscr{L}(\Omega), \mathfrak{e})$ defined by

$$\mathfrak{e}(X) \cdot x = \sup_{P(X)=x} \mathscr{M}(\mathfrak{p}) \cdot P = \sup_{P(X)=x} \inf_{Y \in \mathscr{K}} \mathfrak{p}(Y) \cdot P(Y)$$

for all $X \in \mathscr{L}(\Omega)$ and $x \in \mathbb{R}$, is called the *natural extension* of $\mathfrak{p}$.

The natural extension of a reasonable possibilistic prevision has a number of very interesting and desirable properties, which are summarised in the following theorem.

Theorem 13 (Natural extension theorem)**.** *Let $(\Omega, \mathscr{K}, \mathfrak{p})$ be a reasonable possibilistic prevision, and let $(\Omega, \mathscr{L}(\Omega), \mathfrak{e})$ be its natural extension. Then the following statements hold.*

(*i*) $\mathfrak{e}(X) \leq \mathfrak{p}_v(X)$ *for all* $X \in \mathscr{L}(\Omega)$.
(*ii*) $\mathfrak{e}$ *is a representable possibilistic prevision on* $\mathscr{L}(\Omega)$.
(*iii*) $\mathfrak{e}$ *has greatest representation* $\mathscr{M}(\mathfrak{p})$*, i.e.,* $\mathscr{M}(\mathfrak{e}) = \mathscr{M}(\mathfrak{p})$.
(*iv*) $\mathfrak{e}$ *is point-wise dominated by* $\mathfrak{p}$ *on* $\mathscr{K}$*:* $\mathfrak{e}(X) \leq \mathfrak{p}(X)$ *for all* $X \in \mathscr{K}$.
(*v*) $\mathfrak{e}$ *coincides with* $\mathfrak{p}$ *on* $\mathscr{K}$ *if and only if* $\mathfrak{p}$ *is representable.*
(*vi*) $\mathfrak{e}$ *is the point-wise greatest representable possibilistic prevision on* $\mathscr{L}(\Omega)$ *that is point-wise dominated by* $\mathfrak{p}$ *on* $\mathscr{K}$.
(*vii*) *If* $\mathfrak{p}$ *is representable then* $\mathfrak{e}$ *is the point-wise greatest representable possibilistic prevision on* $\mathscr{L}(\Omega)$ *that coincides with* $\mathfrak{p}$ *on* $\mathscr{K}$.

Proof. Recall that since $\mathfrak{p}$ is assumed to be reasonable, $\mathscr{M}(\mathfrak{p})$ is a normal possibility distribution on $\mathbb{P}$, by Theorem 10(ii). To prove the first statement, observe that by Lemma 42 in Appendix A, $\mathscr{N}(X - x) = \emptyset$ for any real x outside $[\inf X, \sup X]$, whence $\mathfrak{e}(X) \cdot x = \sup\{\mathscr{M}(\mathfrak{p}) \cdot P \colon P \in \mathscr{N}(X - x)\} = \sup \emptyset = 0 = \mathfrak{p}_v(X) \cdot x$. Since for $x \in [\inf X, \sup X]$, $\mathfrak{e}(X) \cdot x \leq 1 = \mathfrak{p}_v(X) \cdot x$, it indeed follows that $\mathfrak{e}(X) \leq \mathfrak{p}_v(X)$. Since $\mathfrak{e}$ is defined as the possibilistic prevision with representation $\mathscr{M}(\mathfrak{p})$, the proof of the second statement is trivial. To prove the fourth statement, consider $X \in \mathscr{K}$ and $x \in \mathbb{R}$. Then

$$\begin{aligned}\mathfrak{e}(X) \cdot x = \sup_{P(X)=x} \mathscr{M}(\mathfrak{p}) \cdot P &= \sup_{P(X)=x} \inf_{Y \in \mathscr{K}} \mathfrak{p}(Y) \cdot P(Y) \\ &\leq \sup_{P(X)=x} \mathfrak{p}(X) \cdot P(X) = \mathfrak{p}(X) \cdot x,\end{aligned}$$

where the inequality holds since $X \in \mathscr{K}$. This tells us that indeed $\mathfrak{e}(X) \leq \mathfrak{p}(X)$. We may use this fact to prove the third statement. Indeed, it follows that for any $P \in \mathbb{P}$,

$$\begin{aligned}\mathscr{M}(\mathfrak{e}) \cdot P = \inf_{X \in \mathscr{L}(\Omega)} \mathfrak{e}(X) \cdot P(X) &\leq \inf_{X \in \mathscr{K}} \mathfrak{e}(X) \cdot P(X) \\ &\leq \inf_{X \in \mathscr{K}} \mathfrak{p}(X) \cdot P(X) = \mathscr{M}(\mathfrak{p}) \cdot P,\end{aligned}$$

which tells us that $\mathscr{M}(\mathfrak{p})$ dominates $\mathscr{M}(\mathfrak{e})$. On the other hand, since $\mathfrak{e}$ has representation $\mathscr{M}(\mathfrak{p})$, its greatest representation $\mathscr{M}(\mathfrak{e})$ dominates $\mathscr{M}(\mathfrak{p})$, whence indeed $\mathscr{M}(\mathfrak{e}) = \mathscr{M}(\mathfrak{p})$. The fifth, sixth and seventh statements are immediate consequences of Theorem 10 and the third statement. □

Let us discuss the various implications of this Natural Extension Theorem for possibilistic previsions. The third statement implies that a reasonable possibilistic prevision and its natural extension 'contain essentially the same information': from the global point of view, they have the same behavioural implications. Statement (vi) tells us that natural extension allows us to correct a reasonable possibilistic prevision into a representable one that is as conservative, or least-committal, as possible. And (vii) shows that natural extension allows us to extend a representable possibilistic prevision to a larger domain, again with minimal behavioural implications.

Anyone familiar with Walley's behavioural theory of imprecise probabilities, will have recognised the striking similarity between this Natural Extension Theorem and Theorems

3.1.2 and 3.4.1 in Walley's book on the subject [41] (see also Theorem 3). Indeed, our Theorem 13 can be formally obtained from the latter simply by making the verbal substitutions listed in Table 1. It appears, therefore, that a possibilistic prevision is a natural generalisation of the notion of a lower prevision. This will become even more apparent in the next section.

lower prevision $\underline{P}$	possibilistic prevision $\mathfrak{p}$
avoiding sure loss	reasonability
coherence	representability
set of linear previsions $\mathscr{M}(\underline{P})$	possibility distribution $\mathscr{M}(\mathfrak{p})$
natural extension $\underline{E}$	natural extension $\mathfrak{e}$
vacuous lower prevision	vacuous possibilistic prevision $\mathfrak{p}_v$
ordering $\leq$ of real numbers	point-wise ordering $\leq$ of possibility distributions
lower envelope	induced possibilistic prevision

TABLE 1. Corresponding notions in the theory of lower previsions (left column) and that of possibilistic previsions (right column)

Example 5 (Continuation of Examples 1 and 4). The natural extension of $(\Omega, \mathscr{K}, \mathfrak{p}_v)$ to $\mathscr{L}(\Omega)$ is the vacuous possibilistic prevision on $\mathscr{L}(\Omega)$. ♦

Example 6 (Continuation of Example 2). It is clear that we can parametrise all linear previsions on $\{h,t\}$ by the value they assume on the event $\{t\}$, i.e., $\mathbb{P} = \{P_\theta \colon \theta \in [0,1]\}$, where for all gambles X on $\{h,t\}$,

$$P_\theta(X) = \theta X(t) + (1-\theta)X(h).$$

I shall therefore in what follows identify $\mathbb{P}$ with $[0,1]$. For the second-order possibility distribution $\mathscr{M}(\mathfrak{p})$ associated with the possibilistic prevision $\mathfrak{p}$, we find that

$$\mathscr{M}(\mathfrak{p}) \cdot P_\theta = \mathfrak{p}(\{t\}) \cdot P_\theta(\{t\}) = f(\theta); \quad \theta \in [0,1].$$

This possibility distribution is normal [$\mathscr{M}(\mathfrak{p}) \cdot P_1 = f(1) = 1$], so $\mathfrak{p}$ is reasonable, and we can consider its natural extension $\mathfrak{e}$, which is given by $\mathfrak{e}(X) \cdot x = \sup_{P_\theta(X)=x} f(\theta)$ for all gambles X on $\{h,t\}$. After some algebraic manipulations, we find that if X is not a constant gamble, then

$$\mathfrak{e}(X) \cdot x = \begin{cases} f\left(\left|\dfrac{x - X(h)}{X(t) - X(h)}\right|\right) & \text{if } \min X \leq x \leq \max X \\ 0 & \text{elsewhere} \end{cases}$$

where of course $\min X = \min\{X(h), X(t)\}$ and similarly for $\max X$. If X is a constant gamble then $\mathfrak{e}(X) \cdot x = 1$ if $x = \min X = \max X$ and zero elsewhere. For $X = I_{\{t\}}$ we find that $\mathfrak{e}(\{t\}) = \mathfrak{p}(\{t\})$, so $\mathfrak{p}$ is representable. ♦

Example 7 (Continuation of Example 3). For the second-order possibility distribution $\mathscr{M}(\mathfrak{p})$ associated with the possibilistic prevision $\mathfrak{p}$, we find that

$$\mathscr{M}(\mathfrak{p}) \cdot P_\theta = \mathfrak{p}(\{t\}) \cdot P_\theta(\{t\}) = g(|2\theta - 1|); \quad \theta \in [0,1].$$

This possibility distribution is normal, so $\mathfrak{p}$ is reasonable, and we can consider its natural extension $\mathfrak{e}$. If the gamble X on $\{h,t\}$ is not constant, then

$$\mathfrak{e}(X) \cdot x = \begin{cases} g\left(\left|\dfrac{x - \dfrac{X(h)+X(t)}{2}}{\dfrac{X(t)-X(h)}{2}}\right|\right) & \text{if } \min X \leq x \leq \max X \\ 0 & \text{elsewhere.} \end{cases}$$

If X is a constant gamble then $\mathfrak{e}(X)\cdot x = 1$ if $x = \min X = \max X$ and zero elsewhere. For $X = I_{\{t\}}$ we find that $\mathfrak{e}(\{t\}) = \mathfrak{p}(\{t\})$, so $\mathfrak{p}$ is representable. Note that $\mathfrak{e}(\{h\}) = \mathfrak{p}(\{t\})$ as well. ⧫

5. An important special case: full possibility distributions

In the previous section, we have studied possibilistic previsions $\mathfrak{p}$ whose values $\mathfrak{p}(X)$ are normal possibility distributions on the set of real numbers. We now turn to the study of the special case where these values are in particular *full* possibility distributions (the notion of fullness will be defined further on).

My reasons for doing so are fourfold. First of all, we shall see that in this special case the analogy between possibilistic and lower previsions becomes very strong. Secondly, and perhaps more importantly, I shall show in Section 8 that this special class of possibilistic previsions can be given a direct behavioural interpretation in terms of so-called buying and price functions, i.e., for this class we can let go of the questionable assumption that there is a true or ideal prevision P_T that the modeller has only limited knowledge of. Thirdly, it will follow from the discussion in Section 6 that any representable possibilistic prevision can be converted into a representable possibilistic prevision whose values are full, and whose behavioural implications are exactly the same, as far as decision-making and statistical reasoning are concerned. And fourthly, full possibility distributions are interesting in their own right, because they are very closely linked with the notion of a fuzzy interval in fuzzy set theory. Thus our study in this section will lead us to the discovery of a surprising connection between the theory of fuzzy intervals and the theory of imprecise probabilities. Another aspect of this connection is uncovered in Section 7.

5.1. **Full possibility distributions and fuzzy intervals.** Consider a map $f\colon D \to [0,1]$, taking values in the real unit interval $[0,1]$, and whose domain D is a convex subset of a topological real vector space. The special cases that we shall consider further on are $D = \mathbb{R}$ and $D = \mathbb{P}$. Then for each $\alpha \in (0,1]$, its *cut set* f_α at level α, or *α-cut*, is defined by

$$f_\alpha = \{x \in D\colon\ f(x) \geq \alpha\}. \tag{8}$$

Notice that the map f is completely determined by its cut sets, since for all x in D

$$f(x) = \sup\{\alpha \in (0,1]\colon x \in f_\alpha\}, \tag{9}$$

and that the cut sets f_α have a special property: for all $\alpha \in (0,1]$,

$$f_\alpha = \bigcap_{0<\gamma<\alpha} f_\gamma. \tag{10}$$

In fact, it is a very easy exercise to show that any family f_α, $\alpha \in (0,1]$ of subsets of D make up the cut sets of some $D \to [0,1]$-map f if and only they satisfy the condition (10), and that this map f is then given by Eq. (9). This seems to have been first noticed by Negoita and Ralescu [38].

A map $f\colon D \to [0,1]$ is called *quasi-concave* if $f(\lambda x + (1-\lambda)y) \geq \min\{f(x), f(y)\}$ for all x and y in D and all λ in $[0,1]$, or equivalently, if its cut sets f_α are convex for all $\alpha \in (0,1]$ [33]. It is called *normal* if $\sup_{x\in D} f(x) = 1$ and *modal* if this supremum is actually achieved, i.e., if $f_1 \neq \emptyset$. Any element of f_1 is called a *modal element* or *mode* of f.

Definition 14. Let D be a convex subset of a topological real vector space. A map $f\colon D \to [0,1]$ is called *full* if its cut sets f_α, $\alpha \in (0,1]$ are convex closed subsets of D, or equivalently, if it is upper semi-continuous and quasi-concave.

Observe that if D is compact, or if the map f has a compact *support* $\{x \in D\colon\ f(x) > 0\}$, then f is normal if and only if it is modal (this is because its cut sets f_α have the finite intersection property). This holds in particular for all full normal second-order possibility distributions, i.e., for the choice $D = \mathbb{P}$ (a compact set).

An other interesting case occurs for the choice $D=\mathbb{R}$. A full normal possibility distribution $f\colon\mathbb{R}\to[0,1]$ with bounded (and therefore compact) support is also modal; for each $\alpha\in(0,1]$, its corresponding cut set f_α is a non-empty bounded and closed (i.e., compact) real interval. In the literature on fuzzy number theory (see for instance [35]), such a map f is also called a (normal) *bounded closed fuzzy interval*, which is so horrible a name that I shall abbreviate it to *fint* in what follows. A fint can also be characterised in an alternative way: it is a $\mathbb{R}-[0,1]$-map with bounded support and at least one mode, that is furthermore non-decreasing and right-continuous to the left of any mode, and non-increasing and left-continuous to the right of any mode.

Two fints f and g can be added: their sum $f+g$ is defined as the fint whose cut sets $(f+g)_\alpha$ are the interval sums $f_\alpha+g_\alpha$ of the cut sets of the terms f and g: $\min(f+g)_\alpha=\min f_\alpha+\min g_\alpha$ and $\max(f+g)_\alpha=\max f_\alpha+\max g_\alpha$. Alternatively, for all real x:

$$(f+g)(x)=\sup_{y\in\mathbb{R}}\min\{f(y),g(x-y)\}.$$

It is also possible to multiply a real number λ with a fint f: the scalar product λf is defined as the fint whose cut sets $(\lambda f)_\alpha$ are the products λf_α of the real number λ and the interval f_α: $\min(\lambda f)_\alpha=\lambda\min f_\alpha$ and $\max(\lambda f)_\alpha=\lambda\max f_\alpha$ if $\lambda\geq0$, and $\min(\lambda f)_\alpha=\lambda\max f_\alpha$ and $\max(\lambda f)_\alpha=\lambda\min f_\alpha$ if $\lambda\leq0$. Alternatively, $(\lambda f)(x)=f(x/\lambda)$ for $\lambda\neq0$, and $0f=I_{\{0\}}$.

Finally, there are two interesting ways of ordering fints. The first is the point-wise order $\leq$, and is defined by $f\leq g$ if and only if $f(x)\leq g(x)$ for all $x\in\mathbb{R}$. The second is an obvious extension of a natural way of ordering intervals by comparing their bounds: $f\preceq g$ if and only if for all $\alpha\in(0,1]$, $\min f_\alpha\leq\min g_\alpha$ and $\max f_\alpha\leq\max g_\alpha$.

5.2. **Full possibilistic previsions and their representability.** We are now ready to start our study of the special class of those possibilistic previsions $\mathfrak{p}$ whose values $\mathfrak{p}(X)$ are fints.

Definition 15. A possibilistic prevision $(\Omega,\mathscr{K},\mathfrak{p})$ is called *full* if it is representable and if its greatest representation $\mathscr{M}(\mathfrak{p})$ is in addition a full $\mathbb{P}-[0,1]$-map.

A representable $\mathfrak{p}$ is therefore full if and only if the cut sets $\mathscr{M}(\mathfrak{p})_\alpha$, $\alpha\in(0,1]$ of its greatest representation $\mathscr{M}(\mathfrak{p})$ are convex weak*-closed subsets of $\mathbb{P}$. They are moreover all non-empty. We shall see further on (see Corollary 18) that the emphasis on the greatest representation in this definition is not necessary: a possibilistic prevision is indeed full if and only if it has *some* full representation.

There is an interesting connection between the fullness of a possibilistic prevision and that of its values.

Theorem 16. *The natural extension of a reasonable possibilistic prevision $(\Omega,\mathscr{K},\mathfrak{p})$ whose values $\mathfrak{p}(X)$, $X\in\mathscr{K}$ are fints, is full. Moreover, a representable possibilistic prevision $(\Omega,\mathscr{K},\mathfrak{p})$ is full if and only if its values $\mathfrak{p}(X)$, $X\in\mathscr{K}$ are fints.*

Proof. Consider a possibilistic prevision $(\Omega,\mathscr{K},\mathfrak{p})$. First, assume that the values $\mathfrak{p}(X)$, $X\in\mathscr{K}$, of $\mathfrak{p}$ are fints. If $\mathfrak{p}$ is reasonable, the possibility distribution $\mathscr{M}(\mathfrak{p})$ on $\mathbb{P}$ is normal. We take a look at its cut sets $\mathscr{M}(\mathfrak{p})_\alpha$ for $\alpha\in(0,1]$:

$$\begin{aligned}\mathscr{M}(\mathfrak{p})_\alpha&=\{P\in\mathbb{P}\colon\mathscr{M}(\mathfrak{p})\cdot P\geq\alpha\}\\&=\{P\in\mathbb{P}\colon\inf_{X\in\mathscr{K}}\mathfrak{p}(X)\cdot P(X)\geq\alpha\}\\&=\bigcap_{X\in\mathscr{K}}\{P\in\mathbb{P}\colon P(X)\in\mathfrak{p}(X)_\alpha\}\\&=\bigcap_{X\in\mathscr{K}}(X^*)^{-1}(\mathfrak{p}(X)_\alpha).\end{aligned}\tag{11}$$

Since, for all $X \in \mathscr{K}$, the evaluation functional X^* is weak*-continuous and linear, and $\mathfrak{p}(X)_\alpha$ is by assumption a closed interval, it follows from Eq. (11) that $\mathscr{M}(\mathfrak{p})_\alpha$ is weak*-closed and convex, so $\mathscr{M}(\mathfrak{p})$ is a full normal (modal!) $\mathbb{P}-[0,1]$-map. Since, by Theorem 13(iii), $\mathscr{M}(\mathfrak{e}) = \mathscr{M}(\mathfrak{p})$ is the greatest representation of $\mathfrak{e}$, $\mathfrak{e}$ is full. If moreover $\mathfrak{p}$ is representable, the same course of reasoning and Theorem 10 tell us that its greatest representation $\mathscr{M}(\mathfrak{p})$ is a full $\mathbb{P}-[0,1]$-map, so $\mathfrak{p}$ itself is full. To complete the proof, it only remains to show that the values of a full possibilistic prevision $(\Omega, \mathscr{K}, \mathfrak{p})$ are fints. Consider $X \in \mathscr{K}$ and $\alpha \in (0,1]$. Then we have to show that $\mathfrak{p}(X)_\alpha$ is a bounded closed interval. From Lemma 17 it follows that since $\mathfrak{p}$ is assumed to have a full greatest representation $\mathscr{M}(\mathfrak{p})$, $\mathfrak{p}(X)_\alpha = X^*(\mathscr{M}(\mathfrak{p})_\alpha)$. Since X^* is weak*-continuous, the direct image $\mathfrak{p}(X)_\alpha$ of the weak*-compact set $\mathscr{M}(\mathfrak{p})_\alpha$ under X^* is a compact subset of the reals. Since X^* is a linear functional, the direct image $\mathfrak{p}(X)_\alpha$ of the convex set $\mathscr{M}(\mathfrak{p})_\alpha$ under X^* is a convex subset of the reals. It follows indeed that $\mathfrak{p}(X)_\alpha$ is a bounded closed real interval. □

Lemma 17. *Let $(\Omega, \mathscr{K}, \mathfrak{p})$ be a representable possibilistic prevision. Let $\pi \colon \mathbb{P} \to [0,1]$ be a representation of $\mathfrak{p}$ that is full. Then for all $\alpha \in (0,1]$ and $X \in \mathscr{K}$, $\mathfrak{p}(X)_\alpha = X^*(\pi_\alpha)$.*

Proof. The proof is based on a compactness argument. Consider arbitrary $X \in \mathscr{K}$, $\alpha \in (0,1]$ and $x \in \mathbb{R}$. If $x \in \mathfrak{p}(X)_\alpha$, then $\mathfrak{p}(X) \cdot x \geq \alpha$, and since π is a representation of $\mathfrak{p}$, $\sup\{\pi(P) \colon P \in \mathscr{N}(X-x)\} \geq \alpha$. Consequently, for any real ε such that $0 < \varepsilon < \alpha$, there is a $P \in \mathscr{N}(X-x)$ such that $\pi(P) \geq \alpha - \varepsilon$, or equivalently, $\mathscr{N}(X-x) \cap \pi_{\alpha-\varepsilon} \neq \emptyset$. Note that since π is full, $\pi_{\alpha-\varepsilon}$ is weak*-closed. Since X^* is weak*-continuous, $\mathscr{N}(X-x) = (X^*)^{-1}(\{x\})$ is weak*-closed as well, and so is the non-empty subset $J_\varepsilon = \mathscr{N}(X-x) \cap \pi_{\alpha-\varepsilon}$ of $\mathbb{P}$. Since the family $(J_\varepsilon \colon 0 < \varepsilon < \alpha)$ of weak*-closed subsets of the weak*-compact set $\mathbb{P}$ is nested, it has the finite intersection property, and it follows that the intersection $\bigcap_{0<\varepsilon<\alpha} J_\varepsilon \neq \emptyset$. Since $\pi_\alpha = \bigcap_{0<\varepsilon<\alpha} \pi_{\alpha-\varepsilon}$, this means that $\mathscr{N}(X-x) \cap \pi_\alpha \neq \emptyset$. Consequently, there is a $P \in \pi_\alpha$ such that $x = P(X) = X^*(P)$, whence $x \in X^*(\pi_\alpha)$. Conversely, if $x \in X^*(\pi_\alpha)$, there is a $P \in \mathscr{N}(X-x)$ such that $\pi(P) \geq \alpha$, whence, since π is a representation of $\mathfrak{p}$, $\mathfrak{p}(X) \cdot x = \sup\{\pi(P) \colon P \in \mathscr{N}(X-x)\} \geq \alpha$, so $x \in \mathfrak{p}(X)_\alpha$. We conclude that indeed $X^*(\pi_\alpha) = \mathfrak{p}(X)_\alpha$. □

Corollary 18. *If a possibilistic prevision $(\Omega, \mathscr{K}, \mathfrak{p})$ has a full representation, then its greatest representation $\mathscr{M}(\mathfrak{p})$ is also full. In other words, a possibilistic prevision is full if and only if it has* some *full representation.*

Proof. Assume that $\mathfrak{p}$ has a representation π that is full. It follows from Lemma 17 that for any $\alpha \in (0,1]$ and for any $X \in \mathscr{K}$, $\mathfrak{p}(X)_\alpha = X^*(\pi_\alpha)$. By assumption, π_α is convex and weak*-compact. Since X^* is weak*-continuous, the continuous direct image $X^*(\pi_\alpha)$ of the weak*-compact set π_α is compact; since X^* is linear, the linear direct image $X^*(\pi_\alpha)$ of the convex set π_α is convex. Consequently, $\mathfrak{p}(X)_\alpha$ is a bounded and closed real interval, so $\mathfrak{p}(X)$ is a fint. In summary, $\mathfrak{p}$ is a representable possibilistic prevision whose values are fints, and is therefore full, by Theorem 16. □

One important and intriguing aspect of possibilistic previsions whose values are fints, is that they are very closely connected with lower and upper previsions. The various faces of this connection will be explored in the rest of this section.

Consider a possibilistic prevision $(\Omega, \mathscr{K}, \mathfrak{p})$ whose values $\mathfrak{p}(X)$, $X \in \mathscr{K}$ are fints. For every $X \in \mathscr{K}$ and $\alpha \in (0,1]$, the cut set $\mathfrak{p}(X)_\alpha$ of $\mathfrak{p}(X)$ is a non-empty bounded closed real interval, and it is therefore completely characterised by its smallest and its greatest elements, which I shall denote by $\underline{\mathfrak{p}}_\alpha(X)$ and $\overline{\mathfrak{p}}_\alpha(X)$ respectively:

$$\underline{\mathfrak{p}}_\alpha(X) = \min\{x \in \mathbb{R} \colon \mathfrak{p}(X) \cdot x \geq \alpha\} = \min \mathfrak{p}(X)_\alpha$$
$$\overline{\mathfrak{p}}_\alpha(X) = \max\{x \in \mathbb{R} \colon \mathfrak{p}(X) \cdot x \geq \alpha\} = \max \mathfrak{p}(X)_\alpha,$$

i.e., $\mathfrak{p}(X)_\alpha = [\underline{\mathfrak{p}}_\alpha(X), \overline{\mathfrak{p}}_\alpha(X)]$. To put this differently, specifying such a possibilistic prevision $(\Omega, \mathscr{K}, \mathfrak{p})$ is completely equivalent to giving a family of pairs of lower and upper previsions $(\Omega,\mathscr{K},\underline{\mathfrak{p}}_\alpha)$ and $(\Omega,\mathscr{K},\overline{\mathfrak{p}}_\alpha)$, $\alpha\in(0,1]$. These will be called the *lower* and *upper cut previsions* of $\mathfrak{p}$. To be sure, there is as yet *nothing* to warrant that these lower and upper previsions will avoid sure loss, be coherent, or even make up conjugate pairs, i.e. that $\underline{\mathfrak{p}}_\alpha(X)=-\overline{\mathfrak{p}}_\alpha(-X)$ whenever both X and $-X$ are in $\mathscr{K}$. The following theorem explores this issue. It will be useful to consider the set $\mathscr{M}(\underline{\mathfrak{p}}_\alpha,\overline{\mathfrak{p}}_\alpha)$ of linear previsions compatible with the pair of lower and upper cut previsions $\underline{\mathfrak{p}}_\alpha$ and $\overline{\mathfrak{p}}_\alpha$:

$$\mathscr{M}(\underline{\mathfrak{p}}_\alpha,\overline{\mathfrak{p}}_\alpha)=\{P\in\mathbb{P}\colon(\forall X\in\mathscr{K})(\underline{\mathfrak{p}}_\alpha(X)\leq P(X)\leq\overline{\mathfrak{p}}_\alpha(X))\}.$$

Theorem 19. *Let $(\Omega,\mathscr{K},\mathfrak{p})$ be a possibilistic prevision whose values $\mathfrak{p}(X)$, $X\in\mathscr{K}$, are fints. Then for all $\alpha\in(0,1]$:*

$$\mathscr{M}(\mathfrak{p})_\alpha=\mathscr{M}(\underline{\mathfrak{p}}_\alpha,\overline{\mathfrak{p}}_\alpha). \tag{12}$$

Moreover, the statements

- (i) *$\mathfrak{p}$ is reasonable;*
- (ii) *its pairs of lower and upper cut previsions $(\underline{\mathfrak{p}}_\alpha,\overline{\mathfrak{p}}_\alpha)$ avoid sure loss for all $\alpha\in(0,1]$;*
- (iii) *$\mathscr{M}(\mathfrak{p})_\alpha=\mathscr{M}(\underline{\mathfrak{p}}_\alpha,\overline{\mathfrak{p}}_\alpha)\neq\emptyset$ for all $\alpha\in(0,1]$;*

are mutually equivalent, and so are the statements:

- $(i)'$ *$\mathfrak{p}$ is representable (and therefore full);*
- $(ii)'$ *its lower and upper cut previsions $(\underline{\mathfrak{p}}_\alpha,\overline{\mathfrak{p}}_\alpha)$ form coherent (and conjugate) pairs of lower and upper previsions, for all $\alpha\in(0,1]$;*
- $(iii)'$ *for all $\alpha\in(0,1]$ and $X\in\mathscr{K}$,*

$$\underline{\mathfrak{p}}_\alpha(X)=\min\{P(X)\colon P\in\mathscr{M}(\underline{\mathfrak{p}}_\alpha,\overline{\mathfrak{p}}_\alpha)\}=\min\{P(X)\colon P\in\mathscr{M}(\mathfrak{p})_\alpha\}$$
$$\overline{\mathfrak{p}}_\alpha(X)=\max\{P(X)\colon P\in\mathscr{M}(\underline{\mathfrak{p}}_\alpha,\overline{\mathfrak{p}}_\alpha)\}=\max\{P(X)\colon P\in\mathscr{M}(\mathfrak{p})_\alpha\}$$

Proof. First of all, observe that for each $\alpha\in(0,1]$

$$\begin{aligned}\mathscr{M}(\mathfrak{p})_\alpha&=\{P\in\mathbb{P}\colon\mathscr{M}(\mathfrak{p})\cdot P\geq\alpha\}\\&=\{P\in\mathbb{P}\colon(\forall X\in\mathscr{K})(\mathfrak{p}(X)\cdot P(X)\geq\alpha)\}\\&=\{P\in\mathbb{P}\colon(\forall X\in\mathscr{K})(\underline{\mathfrak{p}}_\alpha(X)\leq P(X)\leq\overline{\mathfrak{p}}_\alpha(X))\}\\&=\mathscr{M}(\underline{\mathfrak{p}}_\alpha,\overline{\mathfrak{p}}_\alpha),\end{aligned}$$

which indeed proves Eq. (12). We now turn to the first set of equivalences. It is obvious from Theorem 6 that the second and third statements are equivalent. We prove that the first and third statements are equivalent. Assume that $\mathfrak{p}$ is reasonable. It follows from Theorem 10(ii) that $\mathscr{M}(\mathfrak{p})$ is normal. Since, by Theorem 16, the natural extension $\mathfrak{e}$ of $\mathfrak{p}$ is full, and since, by Theorem 13(iii), $\mathscr{M}(\mathfrak{e})=\mathscr{M}(\mathfrak{p})$, it follows that $\mathscr{M}(\mathfrak{p})$ is full. This implies in particular that $\mathscr{M}(\mathfrak{p})_1\neq\emptyset$, whence indeed $\mathscr{M}(\mathfrak{p})_\alpha\supseteq\mathscr{M}(\mathfrak{p})_1\neq\emptyset$ for all $\alpha\in(0,1]$.

Conversely, assume that the sets of compatible linear previsions $\mathscr{M}(\mathfrak{p})_\alpha$, $\alpha\in(0,1]$ are all non-empty. This implies in particular that $\mathscr{M}(\mathfrak{p})$ is modal and therefore normal. Using Theorem 10, we conclude that $\mathfrak{p}$ is reasonable.

We now turn to the second set of equivalences. Again, the equivalence of the second and third statements is obvious. We prove that the first and third statements are equivalent. Assume that $\mathfrak{p}$ is representable (and therefore full, by Theorem 16), and consider $\alpha\in(0,1]$. Since we know that $\mathscr{M}(\mathfrak{p})$ is full and represents $\mathfrak{p}$, it follows from Lemma 17 that for any $X\in\mathscr{K}$, $\{P(X)\colon P\in\mathscr{M}(\mathfrak{p})_\alpha\}=X^*(\mathscr{M}(\mathfrak{p})_\alpha)=\mathfrak{p}(X)_\alpha=[\underline{\mathfrak{p}}_\alpha(X),\overline{\mathfrak{p}}_\alpha(X)]$, which in effect tells us that $(\Omega,\mathscr{K},\underline{\mathfrak{p}}_\alpha)$ is the lower envelope and $(\Omega,\mathscr{K},\overline{\mathfrak{p}}_\alpha)$ the upper envelope of the set of linear previsions $\mathscr{M}(\mathfrak{p})_\alpha$.

Conversely, assume that the third statement holds. We know from Eq. (12) that in general for all $\alpha \in (0,1]$:

$$\begin{aligned}\mathscr{M}(\mathfrak{p})_\alpha &= \{P \in \mathbb{P} \colon (\forall X \in \mathscr{K})(\underline{\mathfrak{p}}_\alpha(X) \leq P(X) \leq \overline{\mathfrak{p}}_\alpha(X))\} \\ &= \bigcap_{X \in \mathscr{K}} (X^*)^{-1}([\underline{\mathfrak{p}}_\alpha(X), \overline{\mathfrak{p}}_\alpha(X)])\end{aligned} \tag{13}$$

Since the map X^* is weak*-continuous and linear, it follows that the $\mathscr{M}(\mathfrak{p})_\alpha$, $\alpha \in (0,1]$ are weak*-closed (and therefore weak*-compact) and convex. Since they are moreover non-empty by assumption, this already tells us that $\mathscr{M}(\mathfrak{p})$ is a normal and full $\mathbb{P}-[0,1]$-map. It remains to show that $\mathscr{M}(\mathfrak{p})$ represents $\mathfrak{p}$. Again, for any X in $\mathscr{K}$ the evaluation functional X^* is weak*-continuous and linear, so the direct image $X^*(\mathscr{M}(\mathfrak{p})_\alpha)$ is a non-empty, compact and convex subset of the reals, or in other words a non-empty bounded closed real interval. This means that

$$X^*(\mathscr{M}(\mathfrak{p})_\alpha) = [\min X^*(\mathscr{M}(\mathfrak{p})_\alpha), \max X^*(\mathscr{M}(\mathfrak{p})_\alpha)] = [\underline{\mathfrak{p}}_\alpha(X), \overline{\mathfrak{p}}_\alpha(X)] = \mathfrak{p}(X)_\alpha,$$

where the second equality follows from the assumption. Consequently, for any $x \in \mathbb{R}$,

$$\begin{aligned}\mathfrak{p}(X)\cdot x &= \sup\{\alpha \in (0,1] \colon x \in \mathfrak{p}(X)_\alpha\} \\ &= \sup\{\alpha \in (0,1] \colon x \in X^*(\mathscr{M}(\mathfrak{p})_\alpha)\} \\ &= \sup\{\alpha \in (0,1] \colon (\exists P \in \mathscr{N}(X-x))(P \in \mathscr{M}(\mathfrak{p})_\alpha)\} \\ &= \sup_{P \in \mathscr{N}(X-x)} \sup\{\alpha \in (0,1] \colon P \in \mathscr{M}(\mathfrak{p})_\alpha\} \\ &= \sup_{P \in \mathscr{N}(X-x)} \mathscr{M}(\mathfrak{p})\cdot P,\end{aligned}$$

which tells us that $\mathfrak{p}$ indeed has representation $\mathscr{M}(\mathfrak{p})$. □

We can take this connection a step further by considering natural extension. Consider a possibilistic prevision $(\Omega, \mathscr{K}, \mathfrak{p})$ whose values are fints, and assume that it is reasonable. Then we know from Theorem 19 that for each $\alpha \in (0,1]$, the pair of lower and upper cut previsions $(\Omega, \mathscr{K}, \underline{\mathfrak{p}}_\alpha)$ and $(\Omega, \mathscr{K}, \overline{\mathfrak{p}}_\alpha)$ avoid sure loss, or equivalently, that the set of linear previsions $\mathscr{M}(\underline{\mathfrak{p}}_\alpha, \overline{\mathfrak{p}}_\alpha)$ is non-empty. We may therefore consider the natural extensions $\underline{E}_\alpha$ and $\overline{E}_\alpha$ of this pair of lower and upper previsions, given by, for all $X \in \mathscr{L}(\Omega)$:

$$\left.\begin{aligned}\underline{E}_\alpha(X) &= \min\{P(X) \colon P \in \mathscr{M}(\underline{\mathfrak{p}}_\alpha, \overline{\mathfrak{p}}_\alpha)\} \\ \overline{E}_\alpha(X) &= \max\{P(X) \colon P \in \mathscr{M}(\underline{\mathfrak{p}}_\alpha, \overline{\mathfrak{p}}_\alpha)\}\end{aligned}\right\}. \tag{14}$$

The following theorem relates these natural extensions to the natural extension $\mathfrak{e}$ of $\mathfrak{p}$ itself.

Theorem 20. *Let $(\Omega, \mathscr{K}, \mathfrak{p})$ be a possibilistic prevision whose values $\mathfrak{p}(X)$, $X \in \mathscr{K}$, are fints. Let $\mathfrak{p}$ be reasonable so that its natural extension $(\Omega, \mathscr{L}(\Omega), \mathfrak{e})$ is defined. Then the lower and upper cut previsions of its natural extension are the natural extensions of its pair of lower and upper cut previsions. In other words, for any $\alpha \in (0,1]$ and any $X \in \mathscr{L}(\Omega)$:*

$$\underline{\mathfrak{e}}_\alpha(X) = \underline{E}_\alpha(X) \quad \textit{and} \quad \overline{\mathfrak{e}}_\alpha(X) = \overline{E}_\alpha(X),$$

where $(\underline{E}_\alpha, \overline{E}_\alpha)$ is the natural extension of the pair of lower and upper cut previsions $(\underline{\mathfrak{p}}_\alpha, \overline{\mathfrak{p}}_\alpha)$, given by Eq. (14).

Proof. Consider $\alpha \in (0,1]$ and $X \in \mathscr{L}(\Omega)$. Since the values of $\mathfrak{p}$ are fints, Eq. (12) holds, so $\mathscr{M}(\mathfrak{p})_\alpha$ is equal to the set $\mathscr{M}(\underline{\mathfrak{p}}_\alpha, \overline{\mathfrak{p}}_\alpha)$ of linear previsions compatible with the lower and upper cut previsions $\underline{\mathfrak{p}}_\alpha$ and $\overline{\mathfrak{p}}_\alpha$. As a result,

$$\left.\begin{aligned}\underline{E}_\alpha(X) &= \min\{P(X) \colon P \in \mathscr{M}(\mathfrak{p})_\alpha\} = \min X^*(\mathscr{M}(\mathfrak{p})_\alpha) \\ \overline{E}_\alpha(X) &= \max\{P(X) \colon P \in \mathscr{M}(\mathfrak{p})_\alpha\} = \max X^*(\mathscr{M}(\mathfrak{p})_\alpha)\end{aligned}\right\}. \tag{15}$$

Since $\mathfrak{p}$ is reasonable, its natural extension $\mathfrak{e}$ is full, by Theorem 16. By the same theorem, $\mathfrak{e}(X)$ is a fint, whence $\mathfrak{e}(X)_\alpha = [\underline{\mathfrak{e}}_\alpha(X), \overline{\mathfrak{e}}_\alpha(X)]$. Since, by Theorem 13, $\mathfrak{e}$ has representation $\mathscr{M}(\mathfrak{e}) = \mathscr{M}(\mathfrak{p})$, Lemma 17 yields, for any $X \in \mathscr{L}(\Omega)$,

$$X^*(\mathscr{M}(\mathfrak{p})_\alpha) = X^*(\mathscr{M}(\mathfrak{e})_\alpha) = \mathfrak{e}(X)_\alpha = [\underline{\mathfrak{e}}_\alpha(X), \overline{\mathfrak{e}}_\alpha(X)], \tag{16}$$

since moreover the cut sets $\mathscr{M}(\mathfrak{p})_\alpha$ and $\mathscr{M}(\mathfrak{e})_\alpha$ of the identical maps $\mathscr{M}(\mathfrak{p})$ and $\mathscr{M}(\mathfrak{e})$ are identical. If we compare Eqs. (15) and (16), we see that indeed $\underline{\mathfrak{e}}_\alpha(X) = \underline{E}_\alpha(X)$ and $\overline{\mathfrak{e}}_\alpha(X) = \overline{E}_\alpha(X)$. □

Example 8 (Continuation of Examples 2 and 6). Observe that $\mathfrak{p}(\{t\})$ is a fint, whose cut sets are given by:

$$\mathfrak{p}(\{t\})_\alpha = [F(\alpha), 1]; \quad \alpha \in (0,1]$$

where F is the so-called *pseudo-inverse* of f, defined by

$$F(\alpha) = \min f_\alpha = \min\{x \in [0,1] \colon f(x) \geq \alpha\}.$$

Consequently, the lower and upper cut previsions of $\mathfrak{p}$ are given by

$$\underline{\mathfrak{p}}_\alpha(\{t\}) = F(\alpha) \text{ and } \overline{\mathfrak{p}}_\alpha(\{t\}) = 1,$$

Similarly, we see that

$$\mathscr{M}(\mathfrak{p})_\alpha = \{P_\theta \colon \theta \geq F(\alpha)\},$$

and this leads to the following expressions for the lower and upper cut previsions of the natural extension $\mathfrak{e}$:[6]

$$\underline{\mathfrak{e}}_\alpha(X) = \min\{P_\theta(X) \colon \theta \geq F(\alpha)\} = \min\{X(t), [1-F(\alpha)]X(h) + F(\alpha)X(t)\}$$

and

$$\overline{\mathfrak{e}}_\alpha(X) = \max\{P_\theta(X) \colon \theta \geq F(\alpha)\} = \max\{X(t), [1-F(\alpha)]X(h) + F(\alpha)X(t)\}.$$

for any gamble X on $\{h,t\}$. In particular we find that $\underline{\mathfrak{e}}_\alpha(\{h\}) = 0$, and $\overline{\mathfrak{e}}_\alpha(\{h\}) = 1 - F(\alpha)$. ♦

We conclude that, loosely speaking, the theory of possibilistic previsions whose values are fints can be recovered by looking at the cuts of the possibilistic previsions and by applying the theory of lower and upper previsions to these cuts. The following theorem can be interpreted to mean that the theory of lower and upper previsions can even be formally embedded into the theory of possibilistic previsions.

Theorem 21. *Let $(\Omega, \mathscr{K}, \mathfrak{p})$ be a possibilistic prevision whose values $\mathfrak{p}(X)$, $X \in \mathscr{K}$, are fints that assume only the values* 0 *and* 1*. In other words, for $X \in \mathscr{K}$ and $x \in \mathbb{R}$:*

$$\mathfrak{p}(X) \cdot x = \begin{cases} 1 & \text{if } \underline{\mathfrak{p}}_1(X) \leq x \leq \overline{\mathfrak{p}}_1(X) \\ 0 & \text{elsewhere}. \end{cases}$$

Then the following statements hold.

- (i) *The greatest representation $\mathscr{M}(\mathfrak{p})$ of $\mathfrak{p}$ assumes only the values* 0 *and* 1*, and $\mathscr{M}(\mathfrak{p})_1$ is the set of linear previsions compatible with the pair of lower and upper cut previsions $(\underline{\mathfrak{p}}_1, \overline{\mathfrak{p}}_1)$.*
- (ii) *The following statements are equivalent: $(\Omega, \mathscr{K}, \mathfrak{p})$ is reasonable; the pair of lower and upper cut previsions $(\underline{\mathfrak{p}}_1, \overline{\mathfrak{p}}_1)$ avoids sure loss; and $\mathscr{M}(\mathfrak{p})_1 \neq \emptyset$.*
- (iii) *The following statements are equivalent: $(\Omega, \mathscr{K}, \mathfrak{p})$ is representable (and therefore full); the pair of lower and upper cut previsions $(\underline{\mathfrak{p}}_1, \overline{\mathfrak{p}}_1)$ is coherent; and $\underline{\mathfrak{p}}_1$ is the lower envelope and $\overline{\mathfrak{p}}_1$ the upper envelope on $\mathscr{K}$ of the set of linear previsions $\mathscr{M}(\mathfrak{p})_1$.*

[6] Observe that this model is a special case of the contamination models discussed in Section 7.

(*iv*) *If $\mathfrak{p}$ is reasonable, then the natural extension $(\Omega,\mathscr{L}(\Omega),\mathfrak{e})$ of $\mathfrak{p}$ assumes only the values* 0 *and* 1*, i.e., for all $X\in\mathscr{L}(\Omega)$ and $x\in\mathbb{R}$,*

$$\mathfrak{e}(X)\cdot x=\begin{cases}1 & \text{if } \underline{\mathfrak{e}}_1(X)\leq x\leq\overline{\mathfrak{e}}_1(X)\\ 0 & \text{elsewhere,}\end{cases}$$

where the pair of lower and upper cut previsions $(\underline{\mathfrak{e}}_1,\overline{\mathfrak{e}}_1)$ is the natural extension of the pair of lower and upper cut previsions $(\underline{\mathfrak{p}}_1,\overline{\mathfrak{p}}_1)$, or in other words, $\underline{\mathfrak{e}}_1$ is the lower envelope and $\overline{\mathfrak{e}}_1$ the upper envelope on $\mathscr{L}(\Omega)$ of the set of linear previsions $\mathscr{M}(\mathfrak{p})_1$.

Proof. We begin with a number of general observations. Since the values of $\mathfrak{p}$ are fints, Eq. (12) holds, which tells us that $\mathscr{M}(\mathfrak{p})_\alpha$ is the set of linear previsions compatible with the pair of lower and upper cut previsions $(\underline{\mathfrak{p}}_\alpha,\overline{\mathfrak{p}}_\alpha)$, $\alpha\in(0,1]$. Since the values of $\mathfrak{p}$ assume only the values 0 and 1, we have that $\underline{\mathfrak{p}}_\alpha=\underline{\mathfrak{p}}_1$, $\overline{\mathfrak{p}}_\alpha=\overline{\mathfrak{p}}_1$ and $\mathscr{M}(\mathfrak{p})_\alpha=\mathscr{M}(\mathfrak{p})_1$, $\alpha\in(0,1]$.

Since by definition $\mathscr{M}(\mathfrak{p})\cdot P=\inf_{X\in\mathscr{K}}\mathfrak{p}(X)\cdot P(X)$, the first statement is now obvious. The second and third statements follow at once from Theorem 19 and the observations above. The same observations together with Theorem 20 prove the fourth statement. □

If we recall that the values of a full possibilistic prevision are fints, and use the addition, scalar multiplication and ordering of fints introduced in Section 5.1, we see yet another way of exposing the analogy between full possibilistic previsions, and coherent lower and upper previsions. The reader is invited to explore the analogy between the theorem below, and the properties of coherent lower previsions listed in Theorem 2.6.1 of [41].

Theorem 22. *Let $(\Omega,\mathscr{K},\mathfrak{p})$ be a full possibilistic prevision. The following properties hold whenever all gambles involved are in $\mathscr{K}$:*

(*i*) $\inf X\preceq\mathfrak{p}(X)\preceq\sup X$;
(*ii*) $(\forall\mu\in\mathbb{R})(\mathfrak{p}(\mu)=\mu)$;
(*iii*) $(\forall\lambda\in\mathbb{R})(\mathfrak{p}(\lambda X)=\lambda\mathfrak{p}(X))$;
(*iv*) $\mathfrak{p}(-X)=-\mathfrak{p}(X)$;
(*v*) $(\forall\mu\in\mathbb{R})(\mathfrak{p}(\mu+X)=\mu+\mathfrak{p}(X)$ *and* $\mathfrak{p}(\mu-X)=\mu-\mathfrak{p}(X))$;[7]
(*vi*) $\mathfrak{p}(X+Y)\leq\mathfrak{p}(X)+\mathfrak{p}(Y)$;
(*vii*) $(\forall\mu\in\mathbb{R})(X\geq Y+\mu\Rightarrow\mathfrak{p}(X)\succeq\mathfrak{p}(Y)+\mu)$.

Proof. Properties (i) to (vi) are immediate reformulations of the properties in Proposition 11, using the special operations for fints introduced above. We prove Property (vii). Consider X and Y in $\mathscr{K}$ and assume that $X\geq Y+\mu$. Consider any $\alpha\in(0,1]$. Since $\mathfrak{p}$ is in particular representable, and has full representation $\mathscr{M}(\mathfrak{p})$, it follows from Lemma 17 that $\mathfrak{p}(X)_\alpha=X^*(\mathscr{M}(\mathfrak{p})_\alpha)$ and $\mathfrak{p}(Y)_\alpha=Y^*(\mathscr{M}(\mathfrak{p})_\alpha)$, whence

$$\begin{aligned}\underline{\mathfrak{p}}_\alpha(X)&=\min\mathfrak{p}(X)_\alpha=\min X^*(\mathscr{M}(\mathfrak{p})_\alpha)=\min\{P(X)\colon P\in\mathscr{M}(\mathfrak{p})_\alpha\}\\&\geq\min\{P(Y+\mu)\colon P\in\mathscr{M}(\mathfrak{p})_\alpha\}=\min\{P(Y)\colon P\in\mathscr{M}(\mathfrak{p})_\alpha\}+\mu\\&=\underline{\mathfrak{p}}_\alpha(Y)+\mu,\end{aligned}$$

and similarly $\overline{\mathfrak{p}}_\alpha(X)\geq\overline{\mathfrak{p}}_\alpha(Y)+\mu$, whence indeed $\mathfrak{p}(X)\succeq\mathfrak{p}(Y)+\mu$. □

Corollary 23. *Let $(\Omega,\mathscr{K},\mathfrak{p})$ be a full possibilistic prevision. If the indicator functions of the events involved belong to $\mathscr{K}$, we have that $\mathfrak{p}(\emptyset)=0$, $\mathfrak{p}(\Omega)=1$, $\mathfrak{p}(\mathrm{co}A)=1-\mathfrak{p}(A)$, and $\mathfrak{p}(A\cup B)\leq\mathfrak{p}(A)+\mathfrak{p}(B)$ if $A\cap B=\emptyset$.*

Remark 3. In Zadeh's work on fuzzy probabilities [47, 48, 49, 52] the so-called fuzzy expectation of a real random variable X on a finite space $\Omega=\{\omega_1,\dots,\omega_n\}$ is defined as the fint:

$$\widetilde{E}(X)=\sum_{k=1}^{n}X(\omega_k)\widetilde{P}(\{\omega_k\}),$$

[7] Observe that, throughout, I identify a real number μ with the fint $I_{\{\mu\}}$.

where the fint $\widetilde{P}(\{\omega_k\})$ is the so-called fuzzy probability of the singleton $\{\omega_k\}$, the summation stands for addition of fints, and the multiplication is scalar multiplication of a real number and a fint.

It follows from this definition that the operator $\widetilde{E}$ is 'additive' in a certain restricted sense: for any *non-negative*[8] X and Y we see that $\widetilde{E}(X+Y) = \widetilde{E}(X) + \widetilde{E}(Y)$, where the addition on the right-hand side is the addition of fints. Similarly, we find that $\widetilde{P}(A \cup B) = \widetilde{P}(A) + \widetilde{P}(B)$ if the events A and B are disjoint, where $\widetilde{P}(A) = \widetilde{E}(I_A)$ is the fuzzy probability of the event A, and similarly for the event B. We see that Zadeh's fuzzy probabilities and expectations are (to a certain extent) *additive*, whereas Theorem 22(v) and Corollary 23 tell us that our possibilistic previsions are only *subadditive*.

Zadeh identifies fuzzy sets, and in particular fuzzy probabilities, with possibility distributions, and he would say for any $x \in \mathbb{R}$ that $\widetilde{P}(\{\omega_k\}) \cdot x$ is the 'possibility that the probability of $\{\omega_k\}$ is x'. Similarly, $\widetilde{E}(X) \cdot x$ is the 'possibility that the expectation of X is x'. So it seems that we could indeed interpret $\widetilde{E}(X)$ as a possibilistic prevision for X, and $\widetilde{E}$ as a possibilistic prevision defined on $\mathscr{L}(\Omega)$. It is then natural to ask whether this possibilistic prevision $\widetilde{E}$ can be representable. Assume that it is. Then we deduce from the previous theorem that $\widetilde{E}(\mu) = \mu$ for any $\mu \in \mathbb{R}$, and therefore in particular for $\mu = I_\Omega = 1$ that $\widetilde{E}(\Omega) = 1$, i.e., $\sum_{k=1}^{n} \widetilde{P}(\{\omega_k\}) = 1$. This can only happen if the fints $\widetilde{P}(\{\omega_k\})$ are actually precise numbers that add to one. This implies that $\widetilde{E}(X)$ is a precise real number as well: there is some $E(X) \in \mathbb{R}$ such that $\widetilde{E}(X) = E(X)$ or equivalently,

$$\widetilde{E}(X) \cdot x = \begin{cases} 1 & \text{if } x = E(X) \\ 0 & \text{if } x \neq E(X). \end{cases}$$

But this implies that the fuzzy expectation $\widetilde{E}$ is actually a precise expectation, and that in particular also the fuzzy probability $\widetilde{P}$ is a precise probability.[9]

This makes us conclude that, unless we want to work with precise rather than fuzzy probabilities, we have to let go either of representability, or of Zadeh's expression for the fuzzy expectation given the atomic fuzzy probabilities $\widetilde{P}(\{\omega_k\})$. le. I have argued before why I think representability is a desirable property, so it should be clear that I don't think this expression makes much sense, and that I believe that it should be replaced (among other things) by the weaker condition of fuzzy subadditivity (see also Theorem 22(v) and Corollary 23). In fact, it seems to me that Zadeh introduces his definition of a fuzzy expectation in an attempt to 'fuzzify' precise probability theory. The results in this section indicate that, in retrospect, he should have tried to fuzzify *imprecise* probability theory instead.

This criticism of fuzzy expectation, and the fuzzy additivity that it entails (to a certain extent) is not new. Dubois and Prade [23] were the first to point out that there may be 'interactivity' between fuzzy probabilities, which makes fuzzy additivity too conservative. They also showed how to correct for this interactivity, in a manner that seems to correspond with the approach I follow here, but is less systematic and less general. A similar critique of fuzzy additivity was given by Halliwell and Shen [32], but it seems they are content with simply replacing fuzzy additivity by fuzzy subadditivity, so the model they suggest need not in general be representable. $\Diamond$

When the domain $\mathscr{K}$ of a possibilistic prevision $\mathfrak{p}$ whose values are fints is a linear subspace of $\mathscr{L}(\Omega)$, it becomes very easy to characterise its representability (and therefore fullness). The result below is a formal counterpart of Theorem 2.5.5 in [41].

[8] This no longer necessarily works if X and Y don't have the same sign in all points of the space Ω.

[9] In a previous version of this paper, this claim was proven in another, more circuitous manner. This earlier proof was incorrect, as Serafín Moral was kind enough to point out to me.

Theorem 24. *Let $\mathscr{K}$ be a linear subspace of $\mathscr{L}(\Omega)$ and let $(\Omega, \mathscr{K}, \mathfrak{p})$ be a possibilistic prevision whose values $\mathfrak{p}(X)$, $X \in \mathscr{K}$, are fints. Then $\mathfrak{p}$ is representable (and therefore full) if and only if it satisfies the axioms*

P1. $(\forall X \in \mathscr{K})(\mathfrak{p}(-X) = -\mathfrak{p}(X))$;
P2. $(\forall X \in \mathscr{K})(\mathfrak{p}(X) \preceq \sup X)$;
P3. $(\forall X \in \mathscr{K})(\forall \lambda > 0)(\mathfrak{p}(\lambda X) = \lambda \mathfrak{p}(X))$;
P4. $(\forall (X,Y) \in \mathscr{K}^2)(\mathfrak{p}(X+Y) \leq \mathfrak{p}(X) + \mathfrak{p}(Y))$.

Proof. It follows at once from Theorem 22 that P1–P4 are necessary. To prove that these conditions are also sufficient, we make use of the lower and upper cut previsions $\underline{\mathfrak{p}}_\alpha$ and $\overline{\mathfrak{p}}_\alpha$, $\alpha \in (0,1]$. If we can prove that they form coherent conjugate pairs of lower and upper previsions, then by Theorem 19 we may conclude that $\mathfrak{p}$ is representable and therefore full. Consider any $\alpha \in (0,1]$. We first prove that $\underline{\mathfrak{p}}_\alpha$ and $\overline{\mathfrak{p}}_\alpha$ form a conjugate pair. For any X in $\mathscr{K}$, $-X$ belongs to $\mathscr{K}$ as well, and we find indeed using P1 that

$$\begin{aligned}\underline{\mathfrak{p}}_\alpha(-X) &= \min\{x \colon \mathfrak{p}(-X)\cdot x \geq \alpha\} \\ &= \min\{x \colon \mathfrak{p}(X)\cdot(-x) \geq \alpha\} \\ &= \min\{-x \colon \mathfrak{p}(X)\cdot x \geq \alpha\} \\ &= -\max\{x \colon \mathfrak{p}(X)\cdot x \geq \alpha\} = -\overline{\mathfrak{p}}_\alpha(X).\end{aligned}$$

It therefore remains to show that, say, $(\Omega, \mathscr{K}, \overline{\mathfrak{p}}_\alpha)$ is a coherent upper prevision. It follows rather easily from P2–P4 that $\overline{\mathfrak{p}}_\alpha$ satisfies

$$\begin{aligned}&(\forall X \in \mathscr{K})(\overline{\mathfrak{p}}_\alpha(X) \leq \sup X) \\ &(\forall X \in \mathscr{K})(\forall \lambda > 0)(\overline{\mathfrak{p}}_\alpha(\lambda X) = \lambda \overline{\mathfrak{p}}_\alpha(X)) \\ &(\forall (X,Y) \in \mathscr{K}^2)(\overline{\mathfrak{p}}_\alpha(X+Y) \leq \overline{\mathfrak{p}}_\alpha(X) + \overline{\mathfrak{p}}_\alpha(Y)),\end{aligned}$$

and a result by Walley [41, Theorem 2.5.5] on the characterisation of coherent upper previsions on linear spaces now tells us that $\overline{\mathfrak{p}}_\alpha$ is indeed coherent. □

6. Decision-making and statistical reasoning

So far, my main concern has been to turn a local possibilistic model into a global one: I have investigated under what conditions a possibilistic prevision can be represented by a second-order possibility distribution. One important reason for doing so, is that such a second-order possibility distribution has certain behavioural implications, which can be represented by a (imprecise) first-order model. This can then in turn be used in decision making, and as an imprecise prior in statistical reasoning. This has been discussed in great detail in a recent paper by Peter Walley [42] (see also [41, Section 5.10] for a more general, if less detailed treatment; and [17]).

Below, I shall only consider decision making, and in its simplest form: what does a modeller's representable possibilistic prevision $\mathfrak{p}$ tell us about whether or not she should buy (or sell) a gamble X for a given price x? Recall that we have assumed that there is some ideal linear prevision P_T, but that the modeller does not know it precisely. If she knew it, however, we may assume that she would buy the gamble for any price $x < P_T(X)$ and sell it for any price $x > P_T(X)$. Based on this observation, we can use coherence arguments to derive from the modeller's uncertainty model $\mathscr{M}(\mathfrak{p})$ about P_T, a supremum acceptable buying price $\underline{P}(X)$ and an infimum acceptable selling price $\overline{P}(X)$ for the gamble X, given by, with obvious notations:

$$\underline{P}(X) = \sup X - \int_{\inf X}^{\sup X} \sup\{\mathscr{M}(\mathfrak{p})\cdot P \colon P(X) \leq x\}\,\mathrm{d}x \tag{17}$$

and

$$\overline{P}(X) = \inf X + \int_{\inf X}^{\sup X} \sup\{\mathscr{M}(\mathfrak{p})\cdot P \colon P(X) \geq x\}\,\mathrm{d}x. \tag{18}$$

The modeller's representable model $\mathfrak{p}$—or its global counterpart $\mathscr{M}(\mathfrak{p})$—implies, through coherence arguments, that she should be willing to buy the gamble X for any price $x < \underline{P}(X)$, and sell it for any price $x > \overline{P}(X)$. Moreover, $\underline{P}(X)$ is the highest supremum acceptable buying price for X that can be so derived from $\mathfrak{p}$, and similarly for $\overline{P}(X)$. In this sense, $\underline{P}(X)$ and $\overline{P}(X)$ summarise the behavioural consequences of the assessments $\mathfrak{p}$ as far as buying and selling the gamble X are concerned. It should be noted that precisely the same type of coherence arguments allow us to turn a hierarchical Bayesian (precise probability) model into a precise first-order linear prevision (see for instance [29, 42]).

Interestingly, if we associate with the gamble X the so-called evaluation functional X^* on $\mathbb{P}$ defined by $X^*(P) = P(X)$ for all P in $\mathbb{P}$, and observe that $\sup X = \sup X^*$ and $\inf X = \inf X^*$ (by Lemma 42 in Appendix A), we see that Eq. (18) can be rewritten as:

$$\begin{aligned}\overline{P}(X) &= \inf X^* + \int_{\inf X^*}^{\sup X^*} \sup_{X^*(P)\geq x} \mathscr{M}(\mathfrak{p})\cdot P\,\mathrm{d}x \\ &= \inf X^* + \int_{\inf X^*}^{\sup X^*} \Pi(\{P\in\mathbb{P}\colon X^*(P)\geq x\})\,\mathrm{d}x = (C)\int_{\mathbb{P}} X^*\,\mathrm{d}\Pi, \quad (19)\end{aligned}$$

and this is the Choquet integral of the real-valued map X^* with respect to the possibility measure Π on $\mathbb{P}$ whose distribution is $\mathscr{M}(\mathfrak{p})$. This gives an alternative interpretation for the so-called *induced first-order upper prevision* $\overline{P}(X)$: it is the upper expectation of the real random variable $X^*(P_T) = P_T(X)$, the unknown ideal prevision for X. Similarly, $\underline{P}(X)$ is the lower expectation of $P_T(X)$.

Example 9 (Continuation of Examples 1,4 and 5). The induced first-order lower and upper previsions of the vacuous possibilistic prevision are the vacuous lower and upper previsions given by

$$\underline{P}_v(X) = \inf X \text{ and } \overline{P}_v(X) = \sup X.$$

They represent minimal behavioural dispositions: according to the behavioural definition of a lower prevision, we should *strictly accept* a gamble X, i.e., accept $X-\varepsilon$ for some $\varepsilon > 0$, if and only if $\underline{P}_v(X) > 0$, or equivalently, if $\inf X > 0$, i.e., if the gamble is sure to yield a strictly positive gain. ♦

To conclude this section, I prove a very interesting alternative formula for computing the induced lower and upper previsions $\underline{P}$ and $\overline{P}$, which will provide a motivation for concentrating on full possibilistic previsions, and lead directly to the connection between full possibilistic previsions and the behavioural model of buying and price functions to be introduced and studied later in the paper. We start from Eq. (19), which expresses $\overline{P}(X)$ as the Choquet integral of the gamble X^* with respect to the possibility measure Π on $\mathbb{P}$ with distribution $\mathscr{M}(\mathfrak{p})$. Using a basic symmetry result, proven by Walley [42] and independently by myself [11, 12, 16], we can rewrite this Choquet integral as

$$\begin{aligned}\overline{P}(X) &= \inf X^* + \int_{\inf X^*}^{\sup X^*} \sup\{\mathscr{M}(\mathfrak{p})\cdot P\colon X^*(P)\geq x\}\,\mathrm{d}x \\ &= \int_0^1 \sup\{X^*(P)\colon \mathscr{M}(\mathfrak{p})\cdot P\geq\alpha\}\,\mathrm{d}\alpha = \int_0^1 \sup\{X^*(P)\colon P\in\mathscr{M}(\mathfrak{p})_\alpha\}\,\mathrm{d}\alpha. \quad (20)\end{aligned}$$

Since the evaluation functional X^* is weak*-continuous and preserves convex combinations, we may actually replace $\mathscr{M}(\mathfrak{p})_\alpha$ by its convex weak*-closure $\overline{\mathrm{co}}(\mathscr{M}(\mathfrak{p})_\alpha)$ in the above expression. Using Proposition 45 in Appendix B, we conclude that $\overline{\mathrm{co}}(\mathscr{M}(\mathfrak{p})_\alpha)$ may in its turn be replaced by $F(\mathscr{M}(\mathfrak{p}))_\alpha$, where $F(\mathscr{M}(\mathfrak{p}))$ is the full closure of the normal possibility distribution $\mathscr{M}(\mathfrak{p})$, i.e., the smallest normal full possibility distribution that dominates $\mathscr{M}(\mathfrak{p})$ (for more details on full closure, see Appendix B). We may conclude that

$$\begin{aligned}\overline{P}(X) &= \int_0^1 \sup\{X^*(P)\colon P \in F(\mathscr{M}(\mathfrak{p}))_\alpha\}\,\mathrm{d}\alpha \\ &= \int_0^1 \sup\{X^*(P)\colon F(\mathscr{M}(\mathfrak{p}))\cdot P \geq \alpha\}\,\mathrm{d}\alpha \\ &= \inf X^* + \int_{\inf X^*}^{\sup X^*} \sup\{F(\mathscr{M}(\mathfrak{p}))\cdot P\colon X^*(P) \geq x\}\,\mathrm{d}x, \qquad (21)\end{aligned}$$

where the last equality follows again from applying the above-mentioned basic symmetry result for Choquet integrals. A similar result holds for the induced lower prevision $\underline{P}(X)$. If we compare Eqs. (19) and (21), we see that $\mathscr{M}(\mathfrak{p})$ and its full closure $F(\mathscr{M}(\mathfrak{p}))$ induce the same upper—and lower—previsions, and *therefore have the same behavioural implications*. This, of course, is one of the reasons why full possibilistic previsions are so important. We summarise these results in the following theorem, which tells us that as far as the behavioural consequences are concerned, we may concentrate on full possibilistic previsions.

Theorem 25. *Let $(\Omega, \mathscr{K}, \mathfrak{p})$ be a representable possibilistic prevision. Then there is a full possibilistic prevision $\mathfrak{p}'$ with the same behavioural implications, i.e., the same induced lower and upper previsions. It is given by*

$$\mathfrak{p}'(X)\cdot x = \sup\{F(\mathscr{M}(\mathfrak{p}))\cdot P\colon P(X) = x\}$$

for all gambles X in $\mathscr{L}(\Omega)$ and all $x \in \mathbb{R}$.

It turns out that if $\mathfrak{p}$ is a full possibilistic prevision, there are still other very useful alternative formulae for calculating the induced lower and upper previsions for a gamble X, involving only the value of the possibilistic prevision $\mathfrak{p}(X)$—a fint—and/or the values of the lower and upper cut previsions $\underline{\mathfrak{p}}_\alpha(X)$ and $\overline{\mathfrak{p}}_\alpha(X)$, $\alpha \in (0,1]$. Indeed, we deduce from Eq. (20) and Theorem 19 that

$$\overline{P}(X) = \int_0^1 \overline{\mathfrak{p}}_\alpha(X)\,\mathrm{d}\alpha$$

and similarly

$$\underline{P}(X) = \int_0^1 \underline{\mathfrak{p}}_\alpha(X)\,\mathrm{d}\alpha.$$

Also, if we observe that for a full possibilistic prevision $(\Omega, \mathscr{K}, \mathfrak{p})$

$$\begin{aligned}\sup\{\mathscr{M}(\mathfrak{p})\cdot P\colon P(X) \geq x\} &= \sup_{y\geq x}\sup\{\mathscr{M}(\mathfrak{p})\cdot P\colon P(X) = y\} = \sup_{y\geq x}\mathfrak{p}(X)\cdot y \\ &= \begin{cases} 1 & \text{if } x \leq \overline{\mathfrak{p}}_1(X) \\ \mathfrak{p}(X)\cdot x & \text{if } x \geq \overline{\mathfrak{p}}_1(X)\end{cases}\end{aligned}$$

for all $X \in \mathscr{K}$ and $x \in \mathbb{R}$, we find that Eq. (18) can be rewritten as

$$\overline{P}(X) = \overline{\mathfrak{p}}_1(X) + \int_{\overline{\mathfrak{p}}_1(X)}^{\sup X} \mathfrak{p}(X)\cdot x\,\mathrm{d}x$$

and similarly

$$\underline{P}(X) = \underline{\mathfrak{p}}_1(X) - \int_{\inf X}^{\underline{\mathfrak{p}}_1(X)} \mathfrak{p}(X)\cdot x\,\mathrm{d}x.$$

Also note that the *imprecision* of the induced first-order model in the gamble X, namely

$$\Delta(X) = \overline{P}(X) - \underline{P}(X) = \int_{\inf X}^{\sup X} \mathfrak{p}(X)\cdot x\,\mathrm{d}x = \int_{-\infty}^{+\infty} \mathfrak{p}(X)\cdot x\,\mathrm{d}x,$$

is the area under the graph of the fint $\mathfrak{p}(X)$!

Example 10 (Continuation of Examples 2, 6 and 8). The induced first-order lower and upper previsions are given by

$$\underline{P}(X) = \min\{X(t), (1-I)X(t) + IX(h)\}$$
$$\overline{P}(X) = \max\{X(t), (1-I)X(t) + IX(h)\}$$

where

$$I = \int_0^1 f(x)\,\mathrm{d}x = 1 - \int_0^1 F(\alpha)\,\mathrm{d}\alpha$$

is the area under the fint $\mathfrak{p}(\{t\})$. The imprecision

$$\Delta(X) = \overline{P}(X) - \underline{P}(X) = I\,|X(t) - X(h)|$$

is proportional to this area. We see that in particular $I = \overline{P}(\{t\}) - \underline{P}(\{t\})$ is the imprecision associated with the given assessment $\mathfrak{p}(\{t\})$.

Since $\underline{P}(\{t\}) = 1 - I$ and $\overline{P}(\{t\}) = 1$, the assessment implies that the modeller should not bet against tails, or on heads, at any non-trivial rate. This is in complete accordance with the initial observation that the available information provides no evidence at all for the occurrence of heads. She should strictly accept bets on tails only at rates strictly smaller that $1 - I$, and this supremum acceptable betting rate decreases as the area I under f becomes larger.

More generally, should the modeller accept a gamble whose reward function is X? Of course, she should accept it if $\min X \geq 0$, as accepting it results in a sure (marginal) gain. She should not accept it if $\max X < 0$, because it is irrational to accept a sure loss. Let us now look at the remaining cases. If $X(t) < 0$ and $X(h) \geq 0$, then $X(t) \leq X(h)$, whence $\underline{P}(X) = X(t) < 0$. This implies that the modeller should not accept a gamble that yields a negative amount of utility for tails. If on the other hand $X(t) > 0$ and $X(h) = -\beta X(t) \leq 0$, were $\beta \geq 0$, then $X(t) \geq X(h)$, whence

$$\underline{P}(X) = (1-I)X(t) + IX(h) = (1 - I - I\beta)X(t),$$

so she should (strictly) accept X only if $\underline{P}(X) > 0$ (because then she is even willing to pay a strictly positive amount in exchange for X) or in other words, if

$$\beta = \left|\frac{X(h)}{X(t)}\right| < \frac{1-I}{I}.$$

We see that the more precise the initial assessment $\mathfrak{p}(\{t\})$, or the smaller I, the higher β can be. The modeller will then accept gambles that lead to a high loss from heads compared to the gain resulting from tails, because she is effectively very sure that the coin will land tails. On the other hand, for very imprecise $\mathfrak{p}(\{t\})$, where I is close to 1, she will only accept gambles where the loss resulting from heads is very small compared to the gain resulting from tails, and her behaviour is now very conservative.

To conclude, let us consider a slightly more complicated decision problem. Suppose that the modeller can take two actions $\mathfrak{a}$ and $\mathfrak{b}$, whose outcome depends on the toss of the coin. The uncertain reward from taking action $\mathfrak{a}$ is then a gamble $X_{\mathfrak{a}}$ on $\{h,t\}$, and similarly, taking action $\mathfrak{b}$ results in an uncertain reward $X_{\mathfrak{b}}$. Which of the two actions should the modeller take, based on the information she has about the coin? Let us only consider the case that $X_{\mathfrak{a}}(h) > X_{\mathfrak{b}}(h)$ and $X_{\mathfrak{b}}(t) > X_{\mathfrak{a}}(t)$. The discussion for the other cases is based on the same principles. The modeller should strictly prefer $\mathfrak{b}$ to $\mathfrak{a}$ if and only if $\underline{P}(X_{\mathfrak{b}} - X_{\mathfrak{a}}) > 0$, because then she is willing to pay a strictly positive amount of utility to exchange the reward associated with $\mathfrak{a}$ for that associated with $\mathfrak{b}$. Since it follows from the assumptions that

$$\underline{P}(X_{\mathfrak{b}} - X_{\mathfrak{a}}) = (1-I)[X_{\mathfrak{b}}(t) - X_{\mathfrak{a}}(t)] + I[X_{\mathfrak{b}}(h) - X_{\mathfrak{a}}(h)]$$

she should striclty prefer $\mathfrak{b}$ to $\mathfrak{a}$ if and only if

$$\frac{X_{\mathfrak{b}}(t)-X_{\mathfrak{a}}(t)}{X_{\mathfrak{a}}(h)-X_{\mathfrak{b}}(h)} > \frac{I}{1-I}.$$

It also follows from the assumptions that

$$\underline{P}(X_{\mathfrak{a}}-X_{\mathfrak{b}}) = X_{\mathfrak{a}}(t)-X_{\mathfrak{b}}(t) < 0,$$

so she should never strictly prefer $\mathfrak{a}$ to $\mathfrak{b}$. There is *indeterminacy*, meaning that there is not enough information in the model $\mathfrak{p}$ to determine which action is to be preferred, when both $\underline{P}(X_{\mathfrak{b}}-X_{\mathfrak{a}}) \leq 0$ and $\underline{P}(X_{\mathfrak{a}}-X_{\mathfrak{b}}) \leq 0$, i.e., if

$$\frac{X_{\mathfrak{b}}(t)-X_{\mathfrak{a}}(t)}{X_{\mathfrak{a}}(h)-X_{\mathfrak{b}}(h)} \leq \frac{I}{1-I}.$$

It is clear that the more imprecise the model is (the closer I is to 1), the greater its indeterminacy. ♦

Example 11 (Continuation of Examples 3 and 7). For all $\alpha \in (0,1]$ we have that $P_\theta \in \mathscr{M}(\mathfrak{p})_\alpha$ if and only if

$$\theta \in \left[0, \frac{1-G(\alpha)}{2}\right] \cup \left[\frac{1+G(\alpha)}{2}, 1\right],$$

where G is the pseudo-inverse of g, defined by $G(\alpha) = \min g_\alpha = \min\{x \in [0,1] \colon g(x) \geq \alpha\}$. We then derive from Eq. (20) that the induced upper prevision $\overline{P}$ and the induced lower prevision $\underline{P}$ are vacuous! It is not difficult to see that the full closure of $\mathscr{M}(\mathfrak{p})$ is the vacuous second-order possibility distribution, and consequently that the full possibilistic prevision that has the same behavioural implications as $\mathfrak{p}$ does, is the vacuous possibilistic prevision. Our model does not allow us to infer non-trivial behavioural dispositions from the linguistic assessment that 'the probability that this coin will land tails on the next toss is close to 1 or close to 0'. We cannot use this assessment to help us in deciding between actions whose outcome depends on the coin toss, which, upon reflection, should not come as too big a surprise. ♦

7. The link between possibilistic previsions and ε-contamination models

In robust statistics, be it frequentist [34] or Bayesian [1, 2], one often encounters so-called contamination (or gross error) models, which are essentially convex mixtures of a precise probability measure (or a linear prevision) with a convex compact set of probability measures (or linear previsions) containing it. Walley has investigated the more general notion of a so-called fuzzy contamination model in his study of second-order possibility distributions [42]. In the present section, I slightly generalise the notion of a (fuzzy) contamination model, and show that this generalisation has an intriguing relationship with triangular and trapezoidal fuzzy numbers, through the mediation of the theory of full possibilistic previsions.

Consider two non-empty sets of linear previsions $\mathscr{Q}_o \subseteq \mathbb{P}$ and $\mathscr{Q}_1 \subseteq \mathbb{P}$. If ε is some number between 0 and 1, then the collection of linear previsions

$$\{(1-\varepsilon)Q_1 + \varepsilon Q_o \colon Q_o \in \mathscr{Q}_o \text{ and } Q_1 \in \mathscr{Q}_1\}$$

is called an *ε-contamination of $\mathscr{Q}_1$ with $\mathscr{Q}_o$*.

Proposition 26. *Let $\mathscr{Q}_o$ and $\mathscr{Q}_1$ be non-empty subsets of $\mathbb{P}$ and let $\varepsilon \in [0,1]$. If $\mathscr{Q}_o$ and $\mathscr{Q}_1$ are convex and weak*-closed, then the ε-contamination*

$$\mathscr{M}_\varepsilon = \{(1-\varepsilon)Q_1 + \varepsilon Q_o \colon Q_o \in \mathscr{Q}_o \text{ and } Q_1 \in \mathscr{Q}_1\}$$

of $\mathscr{Q}_1$ with $\mathscr{Q}_o$ is convex and weak-closed as well.*

Proof. We first show that $\mathscr{M}_\varepsilon$ is convex. Let $\lambda \in [0,1]$ and let P_1 and P_2 be elements of $\mathscr{Q}_\varepsilon$, so there are linear previsions Q_1 and Q_2 in $\mathscr{Q}_1$ and Q_3 and Q_4 in $\mathscr{Q}_o$ such that $P_1 = (1-\varepsilon)Q_1 + \varepsilon Q_3$ and $P_2 = (1-\varepsilon)Q_2 + \varepsilon Q_4$. Then

$$\begin{aligned}\lambda P_1 + (1-\lambda)P_2 &= \lambda[(1-\varepsilon)Q_1 + \varepsilon Q_3] + (1-\lambda)[(1-\varepsilon)Q_2 + \varepsilon Q_4]\\ &= (1-\varepsilon)[\lambda Q_1 + (1-\lambda)Q_2] + \varepsilon[\lambda Q_3 + (1-\lambda)Q_4] \in \mathscr{M}_\varepsilon,\end{aligned}$$

since $\lambda Q_1 + (1-\lambda)Q_2 \in \mathscr{Q}_1$ and $\lambda Q_3 + (1-\lambda)Q_4 \in \mathscr{Q}_o$.

Next, we show that $\mathscr{M}_\varepsilon$ is weak*-closed. Consider the map $f_\varepsilon\colon \mathbb{P}\times\mathbb{P} \to \mathbb{P}$, defined by $f_\varepsilon(P,Q) = (1-\varepsilon)P + \varepsilon Q$. It is obviously weak*-continuous, so the direct image $\mathscr{M}_\varepsilon = f_\varepsilon(\mathscr{Q}_o, \mathscr{Q}_1)$ of the weak*-compact sets $\mathscr{Q}_o$ and $\mathscr{Q}_1$ is a weak*-compact subset of the weak*-closed set $\mathbb{P}$. Since the weak* topology on the dual space $\mathscr{L}(\Omega)^*$ of continuous linear functionals on the linear space $\mathscr{L}(\Omega)$ is Hausdorff, it follows that $\mathscr{Q}_\varepsilon$ is weak*-closed. □

This leads us to the following idea, which generalises Walley's notion of a fuzzy contamination model [42]. Consider, for each $\alpha \in (0,1]$, the following collection of linear previsions:

$$\mathscr{Q}_\alpha = \mathscr{M}_{1-u(\alpha)} = \{u(\alpha)Q_1 + [1-u(\alpha)]Q_o\colon Q_o \in \mathscr{Q}_o \text{ and } Q_1 \in \mathscr{Q}_1\},$$

where u is some transformation of $[0,1]$ with $u(0) = 0$ and $u(1) = 1$. Under fairly nonrestrictive conditions on u, $\mathscr{Q}_o$ and $\mathscr{Q}_1$, these sets of linear previsions $\mathscr{Q}_\alpha$ are the cut sets of some full second-order possibility distribution.

Proposition 27. *Let $\mathscr{Q}_o$ and $\mathscr{Q}_1$ be non-empty convex and weak*-closed subsets of $\mathbb{P}$ such that $\mathscr{Q}_1 \subseteq \mathscr{Q}_o$. Let u be a left-continuous non-decreasing transformation of $[0,1]$, with $u(0) = 0$ and $u(1) = 1$. Then there is a full normal (modal!) possibility distribution π on $\mathbb{P}$ such that for all $\alpha \in (0,1]$:*

$$\pi_\alpha = \mathscr{Q}_\alpha = \{u(\alpha)Q_1 + [1-u(\alpha)]Q_o\colon Q_o \in \mathscr{Q}_o \textit{ and } Q_1 \in \mathscr{Q}_1\}.$$

Proof. It follows from Proposition 26 that the $\mathscr{Q}_\alpha$ are convex and weak*-closed. We show that the $\mathscr{Q}_\alpha$ are nested non-increasingly: we assume that $\alpha \leq \beta$ and show that $\mathscr{Q}_\beta \subseteq \mathscr{Q}_\alpha$. If $u(\alpha) = 1$ then $u(\beta) = 1$ since u is non-decreasing, whence $\mathscr{Q}_\alpha = \mathscr{Q}_\beta$. We may therefore assume that $u(\alpha) < 1$. Let $P \in \mathscr{Q}_\beta$, so there are $Q_o \in \mathscr{Q}_o$ and $Q_1 \in \mathscr{Q}_1$ such that

$$\begin{aligned}P &= u(\beta)Q_1 + [1-u(\beta)]Q_o\\ &= u(\alpha)Q_1 + [u(\beta)-u(\alpha)]Q_1 + [1-u(\beta)]Q_o\\ &= u(\alpha)Q_1 + [1-u(\alpha)]Q_2\end{aligned}$$

where $Q_2 = \gamma Q_1 + (1-\gamma)Q_o$ and $\gamma = [u(\beta)-u(\alpha)]/[1-u(\alpha)]$. Since u is non-decreasing, we have $0 \leq u(\alpha) \leq u(\beta) \leq 1$, so $\gamma \in [0,1]$, and Q_2 is a convex combination of Q_o and Q_1, which both belong to $\mathscr{Q}_o$, since we assumed that $\mathscr{Q}_1 \subseteq \mathscr{Q}_o$. So $Q_2 \in \mathscr{Q}_o$, whence $P \in \mathscr{Q}_\alpha$.

Next, we show that for all $\alpha \in (0,1]$, $\mathscr{Q}_\alpha = \bigcap_{0<\gamma<\alpha} \mathscr{Q}_\gamma$. If we take into account the previous step in this proof, it actually suffices to show that $\mathscr{Q}_\alpha \supseteq \bigcap_{0<\gamma<\alpha} \mathscr{Q}_\gamma$. Assume that $P \in \bigcap_{0<\gamma<\alpha} \mathscr{Q}_\gamma$, so for all $0 < \gamma < \alpha$ there are $Q_\gamma \in \mathscr{Q}_1$ and $R_\gamma \in \mathscr{Q}_o$ such that

$$P = u(\gamma)Q_\gamma + [1-u(\gamma)]R_\gamma. \tag{22}$$

Since, by Tychonov's theorem, the product $\mathscr{Q}_o \times \mathscr{Q}_1$ is compact in the product of the weak*-topology on $\mathscr{L}(\Omega)^*$ with itself, the net $(R_\gamma, Q_\gamma)_{0<\gamma<\alpha}$ in $\mathscr{Q}_o \times \mathscr{Q}_1$ has a subnet that converges to some element (R,Q) of $\mathscr{Q}_o \times \mathscr{Q}_1$. Since u is left-continuous, taking the limit of this subnet in Eq. (22) yields $P = u(\alpha)Q + [1-u(\alpha)]R \in \mathscr{Q}_\alpha$. In conclusion, it follows that the $\mathscr{Q}_\alpha$ are the cut sets π_α of some possibility distribution π. This possibility distribution π is moreover full, since its cut sets are convex and weak*-closed, by Proposition 26. Since $\pi_1 = \mathscr{Q}_1 \neq \emptyset$, π is modal (and therefore also normal). □

For the sake of simplicity, we assume from now on that u is continuous and strictly increasing (and therefore invertible), and that $u(0) = 0$ and $u(1) = 1$. The full possibility distribution whose cut sets π_α are the $\mathscr{Q}_\alpha$ is given by $\pi(P) = \sup\{\alpha \in (0,1]\colon P \in \mathscr{Q}_\alpha\}$, for all $P \in \mathbb{P}$. The lower and upper cut previsions associated with this full possibility distribution are the lower and upper envelopes of its cut sets $\mathscr{Q}_\alpha$, and they are given by, for $\alpha \in (0,1]$:

$$\underline{\mathfrak{p}}_\alpha = u(\alpha)\underline{P}_1 + [1-u(\alpha)]\underline{P}_o$$
$$\overline{\mathfrak{p}}_\alpha = u(\alpha)\overline{P}_1 + [1-u(\alpha)]\overline{P}_o$$

where $\underline{P}_1$ and $\overline{P}_1$ are the lower and upper envelopes of $\mathscr{Q}_1$, and $\underline{P}_o$ and $\overline{P}_o$ are the lower and upper envelopes of $\mathscr{Q}_o$. The full possibilistic prevision $\mathfrak{p}$ induced by this possibility distribution π is given by

$$\mathfrak{p}(X)\cdot x = \begin{cases} 0 & \text{if } x < \underline{P}_o(X) \\ u^{-1}\left(\dfrac{x-\underline{P}_o(X)}{\underline{P}_1(X)-\underline{P}_o(X)}\right) & \text{if } \underline{P}_o(X) \leq x \leq \underline{P}_1(X) \\ 1 & \text{if } \underline{P}_1(X) \leq x \leq \overline{P}_1(X) \\ u^{-1}\left(\dfrac{\overline{P}_o(X)-x}{\overline{P}_o(X)-\overline{P}_1(X)}\right) & \text{if } \overline{P}_1(X) \leq x \leq \overline{P}_o(X) \\ 0 & \text{if } \overline{P}_o(X) < x, \end{cases}$$

where $u^{-1}\colon [0,1] \to [0,1]$ is the inverse of u. The fint $\mathfrak{p}(X)$ is depicted in Figure 1. In the above expression, it is assumed that $\underline{P}_o(X) < \underline{P}_1(X)$ and $\overline{P}_1(X) < \overline{P}_o(X)$. It should be obvious how to proceed if these assumptions are violated. To give one example, if $\underline{P}_o(X) = \underline{P}_1(X)$ and $\overline{P}_1(X) = \overline{P}_o(X)$, we get

$$\mathfrak{p}(X)\cdot x = \begin{cases} 1 & \text{if } \underline{P}_o(X) = \underline{P}_1(X) \leq x \leq \overline{P}_1(X) = \overline{P}_o(X) \\ 0 & \text{elsewhere.} \end{cases}$$

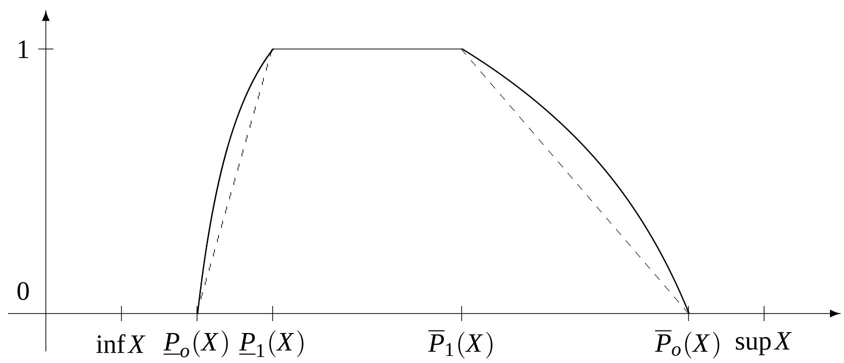

FIGURE 1. The possibilistic prevision $\mathfrak{p}(X)$ of a gamble X derived from a contamination model. The dashed lines depict a trapezoidal fint, and the solid lines the composition of the non-zero part of this fint with the map u^{-1}.

Interestingly, when u, and therefore also u^{-1}, is the identity map on $[0,1]$, then $\mathfrak{p}(X)$ is a so-called trapezoidal fuzzy number (or interval), with support $[\underline{P}_o(X), \overline{P}_o(X)]$ and kernel $[\underline{P}_1(X), \overline{P}_1(X)]$. It is a triangular fuzzy number when $\underline{P}_o(X) = \overline{P}_o(X)$, i.e., when the gamble X has a precise prevision $P_o(X)$. It should be obvious from the preceding discussion that if, conversely, we start out with a collection of trapezodial fuzzy numbers $\mathfrak{p}(X)$, whose supports $[\underline{P}_o(X), \overline{P}_o(X)]$ and kernels $[\underline{P}_1(X), \overline{P}_1(X)]$, $X \in \mathscr{K}$, identify coherent pairs of lower and upper previsions, then the possibilistic prevision is full, and its cut sets are appropriate contaminations of $\mathscr{M}(\underline{P}_1,\overline{P}_1)$ with $\mathscr{M}(\underline{P}_o,\overline{P}_o)$.

8. A BEHAVIOURAL INTERPRETATION: BUYING, SELLING AND PRICE FUNCTIONS

8.1. **Buying and selling functions.** So far, I have started from the assumption that there is uncertainty about the state of the world ω in Ω, that is described by some 'ideal' or 'true' linear prevision P_T. But this assumption is questionable: it does not allow us to give a possibilistic prevision $\mathfrak{p}$ an operational behavioural definition and/or interpretation. Recall that for any gamble X, $\mathfrak{p}(X)\cdot x$ is to be interpreted as a modeller's upper probability for, or the event that the true prevision $P_T(X)$ of X is equal to the real number x, or equivalently, one minus his supremum rate for betting against this event. The problem with this interpretation is that, since P_T is unknown, we cannot determine whether this event occurs, and we cannot therefore 'call the bets'. This implies that the given definition of a possibilistic prevision cannot be made operational.

In a companion paper [17], Peter Walley and I have introduced the notions of *buying*, *selling* and *price functions*, and discussed their more philosphical aspects in some detail (see also the related discussion in [13]). These functions are special hierarchical (second-order) uncertainty models with a clear *behavioural* and *operational* definition and/or interpretation. In this section, I prove a number of mathematical results that are left untouched in that paper, but which have interesting consequences as far as their interpretation is concerned: (i) there is a one-to-one correspondence between a special class of price functions and full possibilistic previsions, and therefore full possibilistic previsions can be made to inherit the behavioural and operational definition of price functions; and (ii) the consistency properties for the (second-order) price functions that were introduced in [17] do not depend on whether we assume that the underlying (first-order) model is a precise or an imprecise probability model. This means that (full) possibilistic previsions can be introduced and justified without assuming the existence of some ideal precise model P_T.

Let me first introduce the basic terminology. We consider two agents. The first is a (male) *subject* who is uncertain about the state ω of the world, where as before Ω denotes the set of all possible such states. Because he is uncertain, he will be disposed to accept certain gambles, and to reject others. Whether he accepts or rejects a gamble can always be checked by simply asking him (this is essentially what makes the model operational). The second (female)[10] agent, called the *modeller*[11], tries to model the subject's uncertainty in the following manner. For any real number x, consider the event $B(X,x)$ that the subject will refuse to buy the gamble X for all prices $y > x$. We denote the modeller's upper probability for this event by $\beta_X(x)$: this number is one minus the modeller's supremum rate for betting against this event. The $\mathbb{R}-[0,1]$-map β_X is called the *buying function* associated with the gamble X. We assume that the modeller represents her beliefs about the subject's behavioural dispositions by specifying buying functions β_X for all gambles X in some subset $\mathscr{K}$ of $\mathscr{L}(\Omega)$.

In the rest of this section, I intend to study buying functions in more detail, and motivate a consistency criterion of *representability* for them. I then show that representable buying functions can be used to define *price functions*, a special subclass of which will turn out to be essentially the same things as the full possibilistic previsions studied earlier.

8.2. **A precise underlying model.** Let us consider the case that the modeller makes the additional assumption that the subject is a Bayesian agent, who has some precise model

[10] In order to distinguish between modeller and agent, I assume that they have different gender. But this is of course only a fictional device of some didactical value, that should not be given any real importance, and should not be driven to extremes. In fact, modeller and agent are the same person in the not unimportant special case where the modeller tries to model her own beliefs, by introspection!

[11] The terminology of 'subject' and 'modeller' allows us to consider many different situations where so-called second-order uncertainty occurs. For a detailed discussion, see [17].

P_T, a linear prevision on Ω, not necessarily known to the modeller. Then she knows that the subject will refuse to buy the gamble X for all prices $y > x$ if and only if $P_T(X) \leq x$. This means that we can represent the event $B(X,x)$ as a subset $B_p(X,x)$ of the set $\mathbb{P}$ of all linear previsions on $\mathscr{L}(\Omega)$, defined as

$$B_p(X,x) = \{P \in \mathbb{P} \colon P(X) \leq x\}.$$

Throughout, I shall use the symbol p to recall that the modeller assumes that the agent's model is a precise, or linear, prevision. If we call

$$\mathscr{B}_p(\mathscr{K}) = \{B_p(X,x) \colon X \in \mathscr{K} \text{ and } x \in \mathbb{R}\},$$

then the collection of buying functions $\{\beta_X : X \in \mathscr{K}\}$ can be modelled as an upper probability $\overline{P}_p$ on the possibility space $\mathbb{P}$, defined on the set of events $\mathscr{B}_p(\mathscr{K})$ as follows:

$$\overline{P}_p(B_p(X,x)) = \beta_X(x), \quad \text{for } X \in \mathscr{K} \text{ and } x \in \mathbb{R}.$$

The upper probability $(\mathbb{P}, \mathscr{B}_p(\mathscr{K}), \overline{P}_p)$ will be called the *p–representing upper probability*, or simply the *p–representation*, of the set of buying functions $\{\beta_X : X \in \mathscr{K}\}$.

Definition 28. A collection $\{\beta_X : X \in \mathscr{K}\}$ of buying functions is called *p–reasonable* if there is some normal possibility measure on $\mathbb{P}$ that is dominated by its *p*–representation $\overline{P}_p$ on its domain $\mathscr{B}_p(\mathscr{K})$. It is called *p–representable* if it can be extended to a normal possibility measure, i.e. if there is some normal possibility measure on $\mathbb{P}$ that coincides with $\overline{P}_p$ on its domain $\mathscr{B}_p(\mathscr{K})$.

In the following theorem, more light is shed on these notions.

Theorem 29. *Let $\{\beta_X : X \in \mathscr{K}\}$ be a collection of buying functions. It is p–reasonable if and only if*

$$\sup_{P \in \mathbb{P}} \inf_{\substack{X \in \mathscr{K}, x \in \mathbb{R} \\ P(X) \leq x}} \beta_X(x) = 1.$$

It is p–representable if and only if it is p–reasonable and if for all $X \in \mathscr{K}$ and $x \in \mathbb{R}$:

$$\beta_X(x) = \sup_{P(X) \leq x} \inf_{Y \in \mathscr{K}} \beta_Y(P(Y)). \tag{23}$$

In that case, the greatest possibility measure Π_p on $\mathbb{P}$ that coincides with the p–representation $\overline{P}_p$ on its domain $\mathscr{B}_p(\mathscr{K})$ has distribution:

$$\pi_p(P) = \inf_{Y \in \mathscr{K}} \beta_Y(P(Y)), \quad P \in \mathbb{P}. \tag{24}$$

Proof. It follows from Theorem 41 in Appendix A that the greatest possibility measure Π_p that is dominated by $\overline{P}_p$ on its domain $\mathscr{B}_p(\mathscr{K})$ has distribution

$$\pi_p(P) = \inf_{\substack{X \in \mathscr{K}, x \in \mathbb{R} \\ P \in B_p(X,x)}} \overline{P}_p(B(X,x)) = \inf_{\substack{X \in \mathscr{K}, x \in \mathbb{R} \\ P(X) \leq x}} \beta_X(x), \tag{25}$$

and that $\overline{P}_p$ dominates some normal possibility measure on its domain if and only if Π_p is normal, that is, if $\Pi_p(\mathbb{P}) = \sup_{P \in \mathbb{P}} \pi_p(P) = 1$. This proves the first part of the theorem. The same Theorem 41 tells us that $\overline{P}_p$ coincides on its domain with some normal possibility measure if and only if Π_p is normal and if Π_p and $\overline{P}_p$ coincide on $\mathscr{B}_p(\mathscr{K})$. The latter means that for all $X \in \mathscr{K}$ and $x \in \mathbb{R}$:

$$\beta_X(x) = \Pi_p(B_p(X,x)) = \sup_{P(X) \leq x} \pi_p(P). \tag{26}$$

Since it follows from this condition that $\beta_X(\cdot)$ is non-decreasing, we find that Eq. (25) may indeed be rewritten as Eq. (24), and that Eq. (26) can be transformed into Eq. (23). Theorem 41 also tells us that if Eq. (23) holds, then π_p is the distribution of the greatest possibility measure that coincides with $\overline{P}_p$ on $\mathscr{B}_p(\mathscr{K})$. $\square$

We can use the possibility distribution π_p to define a buying function on the set $\mathscr{L}(\Omega)$ of all gambles, rather than just on $\mathscr{K}$.

Definition 30. Let $\{\beta_X : X \in \mathscr{K}\}$ be a p–reasonable collection of buying functions. Its p–extension is the collection of buying functions $\{\beta^*_{X,p} : X \in \mathscr{L}(\Omega)\}$, where for any gamble X on Ω and for all x in $\mathbb{R}$:

$$\beta^*_{X,p}(x) = \Pi_p(B_p(X,x)) = \sup_{P(X)\leq x} \pi_p(P) = \sup_{P(X)\leq x} \inf_{\substack{Y\in\mathscr{K}, y\in\mathbb{R}\\ P(Y)\leq y}} \beta_Y(y). \tag{27}$$

It is clear that the model $\{\beta_X : X \in \mathscr{K}\}$, or the associated possibility measure π_p, is a second-order uncertainty model, as it represents the modeller's uncertainty about the subject's uncertainty model. If we make the extra assumption that the modeller would use the subject's model to make decisions, if only she knew what it was, then we can use similar coherence arguments as in Section 6, in [42] and in [41, Section 5.10.5] to derive from the second-order model π_p a first-order lower prevision $\underline{E}^1_p$ that summarises the modeller's dispositions to buy gambles on Ω. For any $X \in \mathscr{L}(\Omega)$, a course of reasoning analogous to the one leading to Eq. (17) tells us that

$$\underline{E}^1_p(X) = \sup X - \int_{\inf X}^{\sup X} \sup\{\pi_p(P) \colon P(X) \leq x\}\,\mathrm{d}x = \sup X - \int_{\inf X}^{\sup X} \beta^*_{X,p}(x)\,\mathrm{d}x.$$

This leads to the following definition.

Definition 31. Let $\{\beta_X : X \in \mathscr{K}\}$ be a p–representable collection of buying functions. Its *first-order p–extension* is the lower prevision $\underline{E}^1_p$ defined for any gamble X on Ω by

$$\underline{E}^1_p(X) = \sup X - \int_{\inf X}^{\sup X} \beta^*_{X,p}(x)\,\mathrm{d}x. \tag{28}$$

$\underline{E}^1_p(X)$ is the modeller's highest supremum acceptable buying price for X that can be derived by coherence arguments from her collection of buying functions $\{\beta_X : X \in \mathscr{K}\}$, under the extra assumption that she believes the subject to have a precise (Bayesian) uncertainty model (a linear prevision), which she would use as her own model for buying gambles, if only she knew what it was.

8.3. **An imprecise underlying model.** Next, let us consider the case that the modeller makes the much weaker additional assumption that the subject is rational in that his uncertainty model is some coherent lower prevision $\underline{P}_T$ on Ω: she does not assume that the subject is a Bayesian agent in the sense that his uncertainty can be modelled by a linear prevision. Then the subject will refuse to buy the gamble X for all prices $y > x$ if and only if $\underline{P}_T(X) \leq x$.[12] This means that we can represent the event $B(X,x)$ as a subset $B_i(X,x)$ of the set $\underline{\mathbb{P}}$ of all coherent lower previsions on $\mathscr{L}(\Omega)$, defined as

$$B_i(X,x) = \{\underline{P} \in \underline{\mathbb{P}} \colon \underline{P}(X) \leq x\}.$$

Throughout, I shall use the symbol i to recall that the modeller assumes that the agent's model is an imprecise probability model, i.e., a coherent lower prevision. If we call

$$\mathscr{B}_i(\mathscr{K}) = \{B_i(X,x) \colon X \in \mathscr{K} \text{ and } x \in \mathbb{R}\},$$

then the collection of buying functions $\{\beta_X : X \in \mathscr{K}\}$ can be modelled as an upper probability $\overline{P}_p$ on the possibility space $\underline{\mathbb{P}}$, defined on the set of events $\mathscr{B}_i(\mathscr{K})$ as follows:

$$\overline{P}_i(B_i(X,x)) = \beta_X(x), \quad \text{for } X \in \mathscr{K} \text{ and } x \in \mathbb{R}.$$

The upper probability $(\underline{\mathbb{P}}, \mathscr{B}_i(\mathscr{K}), \overline{P}_i)$ will be called the *i–representing upper probability*, or simply the *i–representation*, of the set of buying functions $\{\beta_X : X \in \mathscr{K}\}$. For i–representations, similar observations can be made as for p–representations. Let me briefly list the most important ones.

[12] Observe that this means that $\underline{P}_T$ has a so-called *exhaustive* interpretation, see [41, Section 2.3.1] for more details.

Definition 32. A collection $\{\beta_X : X \in \mathscr{K}\}$ of buying functions is called *i–reasonable* if there is some normal possibility measure on $\underline{\mathbb{P}}$ that is dominated by its *i*–representation $\overline{P}_i$ on its domain $\mathscr{B}_i(\mathscr{K})$. It is called *i–representable* if it can be extended to a normal possibility measure, i.e. if there is some normal possibility measure on $\underline{\mathbb{P}}$ that coincides with $\overline{P}_i$ on its domain $\mathscr{B}_i(\mathscr{K})$.

Theorem 33. *Let $\{\beta_X : X \in \mathscr{K}\}$ be a collection of buying functions. It is i–reasonable if and only if*

$$\sup_{\underline{P}\in\underline{\mathbb{P}}} \inf_{\substack{X\in\mathscr{K},x\in\mathbb{R}\\ \underline{P}(X)\leq x}} \beta_X(x) = 1.$$

It is i–representable if and only if it is p–reasonable and if for all $X \in \mathscr{K}$ and $x \in \mathbb{R}$:

$$\beta_X(x) = \sup_{\underline{P}(X)\leq x} \inf_{Y\in\mathscr{K}} \beta_Y(\underline{P}(Y)). \tag{29}$$

In that case, the greatest possibility measure Π_i on $\underline{\mathbb{P}}$ that coincides with the i–representation $\overline{P}_i$ on its domain $\mathscr{B}_i(\mathscr{K})$ has distribution:

$$\pi_i(\underline{P}) = \inf_{Y\in\mathscr{K}} \beta_Y(\underline{P}(Y)), \quad \underline{P} \in \underline{\mathbb{P}}. \tag{30}$$

Definition 34. Let $\{\beta_X : X \in \mathscr{K}\}$ be an *i*–reasonable collection of buying functions. Its *i*–extension is the collection of buying functions $\{\beta^*_{X,i} : X \in \mathscr{L}(\Omega)\}$, where for any gamble X on $\mathscr{K}$ and for all x in $\mathbb{R}$:

$$\beta^*_{X,i}(x) = \Pi_i(B_i(X,x)) = \sup_{\underline{P}(X)\leq x} \pi_i(\underline{P}) = \sup_{\underline{P}(X)\leq x} \inf_{\substack{Y\in\mathscr{K},y\in\mathbb{R}\\ \underline{P}(Y)\leq y}} \beta_Y(y). \tag{31}$$

Definition 35. Let $\{\beta_X : X \in \mathscr{K}\}$ be an *i*–representable collection of buying functions. Its *first-order i–extension* is the lower prevision $\underline{E}^1_i$ defined for any gamble X on Ω by

$$\underline{E}^1_i(X) = \sup X - \int_{\inf X}^{\sup X} \beta^*_{X,i}(x)\,\mathrm{d}x. \tag{32}$$

8.4. **Precision–imprecision equivalence.** Interestingly, and perhaps suprisingly, the notions of reasonability and representability of a collection of buying functions, as well its two types of extensions, do not depend on whether the modeller assumes the subject to have a precise model or not. This is the essence of the following theorem (see [13] for a similar result in another class of hierarchical models).

Theorem 36 (Precision–imprecision equivalence)**.** *Let $\{\beta_X : X \in \mathscr{K}\}$ be a collection of buying functions. Then the following statements hold:*

(i) *$\{\beta_X : X \in \mathscr{K}\}$ is p–reasonable if and only if it is i–reasonable;*
(ii) *$\{\beta_X : X \in \mathscr{K}\}$ is p–representable if and only if it is i–representable;*
(iii) *if $\{\beta_X : X \in \mathscr{K}\}$ is reasonable then its p–extension $\{\beta^*_{X,p} : X \in \mathscr{L}(\Omega)\}$ and its i–extension $\{\beta^*_{X,i} : X \in \mathscr{L}(\Omega)\}$ coincide: $\beta^*_{X,p}(x) = \beta^*_{X,i}(x)$ for all $X \in \mathscr{L}(\Omega)$ and $x \in \mathbb{R}$.*
(iv) *if $\{\beta_X : X \in \mathscr{K}\}$ is representable then its first-order i–extension $\underline{E}^1_i$ and its first-order p–extension $\underline{E}^1_p$ coincide: $\underline{E}^1_i(X) = \underline{E}^1_p(X)$ for all $X \in \mathscr{L}(\Omega)$.*

Proof. We begin by deriving a relation between the distributions π_p on $\mathbb{P}$ and π_i on $\underline{\mathbb{P}}$. For any $\underline{P}$ in $\underline{\mathbb{P}}$, we have, using a course of reasoning similar to the one leading to Eq. (25):

$$\pi_i(\underline{P}) = \inf_{\substack{X\in\mathscr{K},x\in\mathbb{R}\\ \underline{P}(X)\leq x}} \beta_X(x) = \inf_{P\in\mathscr{M}(\underline{P})} \inf_{\substack{X\in\mathscr{K},x\in\mathbb{R}\\ P(X)\leq x}} \beta_X(x)$$

since for all $X \in \mathscr{K}$, $\underline{P}(X) = \min\{P(X) : P \in \mathscr{M}(\underline{P})\}$ by Theorem 6, and therefore $\underline{P}(X) \leq x$ if and only if $(\exists P \in \mathscr{M}(\underline{P}))(P(X) \leq x)$. If we take into account Eq. (25), this leads to

$$\pi_i(\underline{P}) = \inf_{P\in\mathscr{M}(\underline{P})} \pi_p(P). \tag{33}$$

Next, consider any $X \in \mathscr{L}(\Omega)$ and $x \in \mathbb{R}$. We want to prove that

$$\Pi_i(B_i(X,x)) = \sup_{\underline{P}(X)\leq x} \pi_i(\underline{P}) = \sup_{P(X)\leq X} \pi_p(P) = \Pi_p(B_p(X,x)). \tag{34}$$

When $x < \inf X$, we have that both $B_i(X,x) = \emptyset$ and $B_p(X,x) = \emptyset$, by Theorems 2 and 5, so the expressions on both sides are zero and therefore equal. Assume therefore that $x \geq \inf X$. Since for any $P \in \mathbb{P}$, we have that $P \in \underline{\mathbb{P}}$ and $\mathscr{M}(P) = \{P\}$, we may deduce from Eq. (33) that $\pi_i(P) = \pi_p(P)$, whence

$$\sup_{\underline{P}(X)\leq x} \pi_i(\underline{P}) \geq \sup_{P(X)\leq x} \pi_i(P) = \sup_{P(X)\leq x} \pi_p(P).$$

To prove the converse inequality, consider $\underline{P} \in \underline{\mathbb{P}}$ such that $\underline{P}(X) \leq x$ [this is always possible since $x \geq \inf X$]. By Theorem 6, there is some $Q \in \mathscr{M}(\underline{P})$ such that $Q(X) = \underline{P}(X) \leq x$ and consequently, taking into account Eq. (33),

$$\pi_i(\underline{P}) = \inf_{P\in\mathscr{M}(\underline{P})} \pi_p(P) \leq \pi_p(Q) \leq \sup_{P(X)\leq x} \pi_p(P),$$

whence indeed also

$$\sup_{\underline{P}(X)\leq x} \pi_i(\underline{P}) \leq \sup_{P(X)\leq x} \pi_p(P).$$

The rest of the proof is now fairly straightforward. To prove the first statement, choose any $X \in \mathscr{L}(\Omega)$ and let $x \geq \sup X$. Then since for any coherent lower prevision $\underline{P}$ and for any linear prevision P, we have that $\underline{P}(X) \leq \sup X$ and $P(X) \leq \sup X$, it follows that $B_i(X,x) = \underline{\mathbb{P}}$, $B_p(X,x) = \mathbb{P}$, and Eq. (34) becomes

$$\Pi_i(\underline{\mathbb{P}}) = \sup_{\underline{P}\in\underline{\mathbb{P}}} \pi_i(\underline{P}) = \sup_{P\in\mathbb{P}} \pi_p(P) = \Pi_p(\mathbb{P}).$$

Since it follows from the proof of Theorem 29 that $\{\beta_X : X \in \mathscr{K}\}$ is p–reasonable if and only if $\Pi_p(\mathbb{P}) = 1$, and similarly, that it is i–reasonable if and only if $\Pi_i(\underline{\mathbb{P}}) = 1$, this completes the proof for the first statement. To prove the second statement, recall from Theorem 29 that $\{\beta_X : X \in \mathscr{K}\}$ is p–representable if and only if

$$\beta_X(x) = \Pi_p(B_p(X,x)) \text{ for all } X \in \mathscr{K} \text{ and } x \in \mathbb{R},$$

which using Eq. (34) is equivalent to

$$\beta_X(x) = \Pi_i(B_i(X,x)) \text{ for all } X \in \mathscr{K} \text{ and } x \in \mathbb{R},$$

which is in turn equivalent to the i–representability of $\{\beta_X : X \in \mathscr{K}\}$. To prove the third statement, combine Eq. (34) with Eqs. (27) and (31). The fourth statement is an immediate consequence for the third, taking into account Eqs. (28) and (32). □

As an immediate result of this theorem, we can omit reference to whether the subject's model is assumed to be precise or imprecise: we shall simply speak about the reasonability of a collection $\{\beta_X : X \in \mathscr{K}\}$ of buying functions, its representability, its extension $\{\beta_X^* : X \in \mathscr{L}(\Omega)\}$ and its first-order extension $\underline{E}^1$.

8.5. **Properties of representable buying functions.** From now on, we concentrate on collections of buying functions that are representable. A number of their most important properties are listed in the following theorem.

Theorem 37. *Let $\{\beta_X : X \in \mathscr{K}\}$ be a representable collection of buying functions. Then for all real numbers x, y and c, and for all gambles X, Y in $\mathscr{K}$:*

B1. $0 \leq \beta_X(x) \leq 1$;
B2. *if* $x > \sup X$ *then* $\beta_X(x) = 1$;
B3. *if* $x < \inf X$ *then* $\beta_X(x) = 0$;
B4. *if* $c > 0$ *and* $cX \in \mathscr{K}$ *then* $\beta_{cX}(cx) = \beta_X(x)$;
B5. *if* $X+Y \in \mathscr{K}$ *then* $\beta_{X+Y}(x+y) \leq \max\{\beta_X(x), \beta_Y(y)\}$;
B6. β_X *is a non-decreasing function;*

B7. *if* $X+c\in\mathscr{K}$ *then* $\beta_{X+c}(x+c)=\beta_X(x)$*;*
B8. *if the constant gamble* c *belongs to* $\mathscr{K}$*, then* $\beta_c(x)=1$ *if* $x\geq c$ *and* $\beta_c(x)=0$ *if* $x<c$*.*
B9. *if* $-X\in\mathscr{K}$ *then* $\max\{\beta_X(x),\beta_{-X}(-x)\}=1$*;*

Proof. The proof of B1 is obvious. To prove B2–B5, we work with the p–representability of $\{\beta_X\colon X\in\mathscr{K}\}$, and use Theorem 29, and the notation introduced there. If $x>\sup X$ then $P(X)\leq\sup X<x$ for all $P\in\mathbb{P}$, whence, by Eqs. (23) and (24), $\beta(x)=\sup_{P\in\mathbb{P}}\pi_p(P)=1$ [recall that π_p is normal because $\{\beta_X\colon X\in\mathscr{K}\}$ is in particular p–reasonable]. This proves B2. To prove B3, recall that $P(X)\geq\inf X$ for all $P\in\mathbb{P}$, so if $x<\inf X$ there is no P in $\mathbb{P}$ such that $P(X)\leq x$. By Eqs. (23) and (24), $\beta_X(x)=\sup\emptyset=0$. To see that B4 holds, use Eq. (23) and take into account that for any $P\in\mathbb{P}$, $P(cX)=cP(X)$, so $P(cX)\leq cx$ if and only if $P(X)\leq x$. To prove B5, observe that

$$\begin{aligned}\{P\in\mathbb{P}\colon P(X+Y)\leq x+y\}&=\{P\in\mathbb{P}\colon P(X)+P(Y)\leq x+y\}\\&\subseteq\{P\in\mathbb{P}\colon P(X)\leq x\}\cup\{P\in\mathbb{P}\colon P(Y)\leq y\},\end{aligned}$$

and use Eq. (23). To prove B6, assume that $x\leq y$. Then $\{P\in\mathbb{P}\colon P(X)\leq x\}\subseteq\{P\in\mathbb{P}\colon P(X)\leq y\}$, whence $\beta_X(x)\leq\beta_Y(y)$, using Eq. (23). To prove B7, recall that $P(X+c)=P(X)+c$ for all $P\in\mathbb{P}$, and again use Eq. (23). B8 is proven using Eq. (23), and the observation that $P(c)=c$ for all $P\in\mathbb{P}$. Finally, B9 follows at once from Eq. (23) and

$$\mathbb{P}=\{P\in\mathbb{P}\colon P(X)\leq x\}\cup\{P\in\mathbb{P}\colon P(-X)\leq -x\}.\quad\square$$

8.6. **Price functions.** If we observe that selling a gamble X for price x is the same thing as bying the gamble $-X$ for price $-x$, it is obvious that $\beta_{-X}(-x)$ is the modeller's upper probability for the event that the subject will refuse to sell the gamble X for all prices $y<x$. I also denote $\beta_{-X}(-x)$ as $\sigma_X(x)$ and call σ_X the modeller's (conjugate) *selling function* for X.

We can use representable β_X and σ_X to define the modeller's *price function* ρ_X for X as follows:

$$\rho_X(x)=\min\{\beta_X(x),\sigma_X(x)\}=\min\{\beta_X(x),\beta_{-X}(-x)\},\quad x\in\mathbb{R}$$

It is very easy to derive the following interesting properties of price functions from Theorem 37.

Proposition 38. *Let* $\{\beta_X\colon X\in\mathscr{K}\}$ *be a representable collection of buying functions, where* $\mathscr{K}=-\mathscr{K}$*.*[13] *Consider any* X *in* $\mathscr{K}$*. If the buying function* β_X *is right-continuous, the associated price function* ρ_X *is a fint, or in other words, a fuzzy number with compact support* $\{x\in\mathbb{R}\colon\rho_X(x)>0\}$*, with at least one mode* m_X *for which* $\rho_X(m_X)=1$*, and such that* ρ_X *is non-decreasing and right-continuous to the left of* m_X*, and non-increasing and left-continuous to the right of* m_X*.*

Also observe that for all $x\in\mathbb{R}$ and any mode m_X

$$\rho_X(x)=\begin{cases}\beta_X(x) & \text{if } x\leq m_X\\ \sigma_X(x) & \text{if } x\geq m_X,\end{cases}$$

and that

$$\beta_X(x)=\begin{cases}\rho_X(x) & \text{if } x\leq m_X\\ 1 & \text{if } x\geq m_X\end{cases}\quad\text{and}\quad\sigma_X(x)=\begin{cases}1 & \text{if } x\leq m_X\\ \rho_X(x) & \text{if } x\geq m_X.\end{cases}$$

To the left of any mode m_X, ρ_X describes the modeller's beliefs about the subject's dispositions to buy X, whereas her beliefs about the subject's dispositions to sell X are summarised by the values of ρ_X to the right of m_X.

[13] Assuming that $\mathscr{K}=-\mathscr{K}$ allows us to concentrate on buying functions alone, and forget about selling functions. Of course, this assumption is not necessary if conjugate buying and selling functions are both defined on $\mathscr{K}$.

8.7. **The connection with possibilistic previsions.** We are now ready to uncover, in the following two theorems, the very close connection that exists between a special class of price functions on the one hand, and full possibilistic previsions on the other.

Theorem 39. *Let $\{\beta_X : X \in \mathscr{K}\}$ be a representable collection of right-continuous buying functions, where $\mathscr{K} \subseteq \mathscr{L}(\Omega)$ is a set of gambles on Ω such that $-\mathscr{K} = \mathscr{K}$. Let $\{\rho_X : X \in \mathscr{K}\}$ be the associated collection of price functions. Then the possibilistic prevision $(\Omega, \mathscr{K}, \mathfrak{p})$ defined by*

$$\mathfrak{p}(X) \cdot x = \rho_X(x) = \min\{\beta_X(x), \sigma_X(x)\} = \min\{\beta_X(x), \beta_{-X}(-x)\}$$

for all $X \in \mathscr{K}$ and $x \in \mathbb{R}$, is full. Moreover, its greatest representation $\mathscr{M}(\mathfrak{p})$ is identical to the distribution π_p of the greatest possibility measure that coincides with the p-representation of $\{\beta_X : X \in \mathscr{K}\}$ on its domain $\mathscr{B}_p(\mathscr{K})$: for all $P \in \mathbb{P}$,

$$\mathscr{M}(\mathfrak{p}) \cdot P = \inf_{Y \in \mathscr{K}} \rho_Y(P(Y)) = \inf_{Y \in \mathscr{K}} \beta_Y(P(Y)) = \pi_p(P). \tag{35}$$

Proof. We deduce from Proposition 38 that the values $\mathfrak{p}(X) = \rho_X$ of the possibilistic prevision $\mathfrak{p}$ are fints. Moreover, for any P in $\mathbb{P}$ we find that,

$$\begin{aligned}\mathscr{M}(\mathfrak{p}) \cdot P = \inf_{Y \in \mathscr{K}} \mathfrak{p}(Y) \cdot P(Y) &= \inf_{Y \in \mathscr{K}} \min\{\beta_Y(P(Y)), \beta_{-Y}(-P(Y))\} \\ &= \min\left\{\inf_{Y \in \mathscr{K}} \beta_Y(P(Y)), \inf_{Y \in \mathscr{K}} \beta_{-Y}(P(-Y))\right\} = \pi_p(P),\end{aligned}$$

since $-\mathscr{K} = \mathscr{K}$, and this proves Eq. (35). If we can show that $\mathfrak{p}$ is representable, the proof is complete, by Theorem 16. Now, since the collection of buying functions $\{\beta_X : X \in \mathscr{K}\}$ is representable, and therefore p–representable, we deduce from Theorem 29 and Eq. (35) that for all $X \in \mathscr{K}$ and $x \in \mathbb{R}$

$$\beta_X(x) = \sup_{P(X) \le x} \pi_p(P) = \sup_{P(X) \le x} \mathscr{M}(\mathfrak{p}) \cdot P = \sup_{y \le x} \sup_{P(X) = y} \mathscr{M}(\mathfrak{p}) \cdot P = \sup_{y \le x} \mathfrak{e}(X) \cdot y$$

where $\mathfrak{e}$ is the natural extension of $\mathfrak{p}$, which is well-defined since $\mathscr{M}(\mathfrak{p}) = \pi_p$ is normal because $\{\beta_X : X \in \mathscr{K}\}$ is assumed to be representable. In a similar manner, we find that

$$\sigma_X(x) = \beta_{-X}(-x) = \sup_{y \le -x} \mathfrak{e}(-X) \cdot y = \sup_{y \ge x} \mathfrak{e}(-X) \cdot -y = \sup_{y \ge x} \mathfrak{e}(X) \cdot y$$

using Proposition 11 and the fact that $\mathfrak{e}$ is representable on $\mathscr{L}(\Omega)$ [by Theorem 13]. Recall that $\mathfrak{e}(X)$ is a fint [by Theorem 16], and let m_X be any mode of $\mathfrak{e}(X)$. Since a fint is non-decreasing to the left of a mode, and non-increasing to the right, we find that $\beta_X(x) = \mathfrak{e}(X) \cdot x$ and $\sigma_X(x) = 1$ for $x \le m_X$, whereas $\beta_X(x) = 1$ and $\sigma_X(x) = \mathfrak{e}(X) \cdot x$ for $x \ge m_X$. In both cases, $\mathfrak{p}(X) \cdot x = \min\{\beta_X(x), \sigma_X(x)\} = \mathfrak{e}(X) \cdot x$, which tells us that $\mathfrak{p}$ is representable, by Theorem 13. □

Theorem 40. *Let $(\Omega, \mathscr{K}, \mathfrak{p})$ be a full possibilistic prevision, and assume that $-\mathscr{K} = \mathscr{K}$. Then the collection of buying functions $\{\beta_X : X \in \mathscr{K}\}$ defined by*

$$\beta_X(x) = \begin{cases} \mathfrak{p}(X) \cdot x & \text{if } x \le m_X \\ 1 & \text{if } x \ge m_X, \end{cases} \tag{36}$$

where m_X is any mode of $\mathfrak{p}(X)$, is representable. Moreover, the distribution π_p of the greatest possibility measure that coincides with the p-representation of $\{\beta_X : X \in \mathscr{K}\}$, is identical to the greatest representation $\mathscr{M}(\mathfrak{p})$ of $\mathfrak{p}$: for all $P \in \mathbb{P}$,

$$\pi_p(P) = \inf_{Y \in \mathscr{K}} \beta_Y(P(Y)) = \inf_{Y \in \mathscr{K}} \mathfrak{p}(Y) \cdot P(Y) = \mathscr{M}(\mathfrak{p}) \cdot P. \tag{37}$$

Proof. First of all, it follows from Eq. (36) and Proposition 11 that for all $X \in \mathscr{K}$ and $x \in \mathbb{R}$, $\mathfrak{p}(X) \cdot x = \min\{\beta_X(x), \beta_{-X}(-x)\}$. Consequently, since also $-\mathscr{K} = \mathscr{K}$, we may write for all $P \in \mathbb{P}$ that

$$\pi_p(P) = \inf_{Y \in \mathscr{K}} \beta_Y(P(Y)) = \min\left\{\inf_{Y \in \mathscr{K}} \beta_Y(P(Y)), \inf_{Y \in \mathscr{K}} \beta_{-Y}(P(-Y))\right\}$$
$$= \inf_{Y \in \mathscr{K}} \min\{\beta_Y(P(Y)), \beta_{-Y}(-P(Y))\} = \inf_{Y \in \mathscr{K}} \mathfrak{p}(X) \cdot P(Y) = \mathscr{M}(\mathfrak{p}) \cdot P,$$

which tells us that Eq. (37) holds. We must still prove that the collection of buying functions $\{\beta_X : X \in \mathscr{K}\}$ is representable. We shall look at the p–representability, and use Theorem 29. Consider X in $\mathscr{K}$. Since $\mathfrak{p}(X)$ is a fint and therefore non-decreasing to the left of a mode m_X, we have for any $x \le m_X$ that

$$\beta_X(x) = \mathfrak{p}(X) \cdot x = \sup_{y \le x} \mathfrak{p}(X) \cdot y$$

and since $\mathfrak{p}$ is representable, and $\mathscr{M}(\mathfrak{p}) = \pi_p$,

$$= \sup_{y \le x} \sup_{P(X)=y} \mathscr{M}(\mathfrak{p}) \cdot P = \sup_{y \le x} \sup_{P(X)=y} \pi_p(P) = \sup_{P(X) \le x} \pi_p(P).$$

For $x \ge m_X$ we find in a similar way that

$$\beta_X(x) = 1 = \mathfrak{p}(X) \cdot m_X = \sup_{y \le x} \mathfrak{p}(X) \cdot x = \sup_{P(X) \le x} \pi_p(P).$$

By Theorem 29, this means that the collection of buying functions $\{\beta_X : X \in \mathscr{K}\}$ is representable. □

Remark 4. In the companion paper [17], Peter Walley and I only considered the definition of a collection of buying functions $\{\beta_X : X \in \mathscr{K}\}$ where $\mathscr{K}$ was some linear subspace of $\mathscr{L}(\Omega)$. We called such a collection *pos–coherent* if it satisfies Properties B1-B5 in Theorem 37, together with the right-continuity of all the β_X. We showed that such a pos-coherent collection of buying functions is representable in the sense defined above. Theorems 39 and 40 tell us that the present discussion is more general, in that more general collections of buying functions are considered, but that essentially the same representation results can be derived in this more general context. ♢

9. Conclusion

I have presented a behavioural model—possibilistic previsions—for representing and making inferences from vague probability assessments, and I have indicated that there are ways of making this model operational, through the definition of buying and price functions. Possibilistic previsions are formally closely linked to fuzzy probabilities, but they have different interpretations, and they are governed by a different calculus. I have argued in several places why I believe the present model to be superior to that of fuzzy probabilities.

The starting point for introducing this model is that in many cases, linguistic information can be modelled by a possibility measure (or distribution), a point that has been argued in the context of the behavioural theory of imprecise probabilities in a fairly recent paper by Peter Walley and myself [43]. In the same paper, we warn that there may be specific situations where possibility measures do not capture all the information conveyed by statements in natural language. This implies that the present model may not be able to capture and deal with all types of vague probability assessments. But I feel that it will be useful in quite a number of practical situations.

It also deserves to be mentioned here that the theory of possibilistic previsions generalises possibilistic logic [22], in very much the same way as the theory of imprecise probabilities extends classical propositional logic [10, 14]. Let me sketch briefly how such a generalisation comes about. In possibilistic logic there is a representation in terms of a possibility distribution on sets of worlds. It is well-known that worlds can be seen as ultrafilters of propositions, and these are in one-to-one correspondence with zero-one-valued probability measures (see, for instance, [41, Section 2.9.8] and [10, 14]). The generalisation of possibilistic logic then essentially consists in our considering possibility distributions on all linear previsions rather than just zero-one-valued probability measures.

Finally, I want to point out that, although the present model in its full generality seems fairly complicated, I know of at least a few cases where it is computationally tractable and even efficient to work with. One example is that of the so-called *possibilistic (probability) mass functions*, where Ω is finite, and the possibilistic prevision $\mathfrak{p}$ is defined on the singleton sets of Ω (whence the 'mass function'). More details about this fairly simple special case are given in [28].

ACKNOWLEDGEMENTS

While doing the research for this paper, I was subsequently a Postdoctoral Fellow and a Research Associate of the Flemish Fund for Scientific Research – Flanders (FWO). Writing this paper was partially supported by research grant G.0139.01 of the FWO.

Most of the research was carried out during my one-year stay as a Visiting Research Scholar at Binghamton University in 1997. For the first part of the paper, I have benefitted from discussions with George Klir, Yin Pan and Leen Gilbert. The second part of the paper, dealing with the behavioural interpretation of possibilistic previsions, results from a close collaboration with Peter Walley. Unfortunately, foreseeing he would not have the time to help me write the paper, he chose to withdraw as a second author.

I would like to dedicate this paper to Heidi, for being the wonderful woman that she is.

APPENDIX A. AUXILIARY RESULTS

Theorem 41. *(Possibilistic extension theorem) Let X be a non-empty set and $\mathscr{A}$ a non-empty collection of subsets of X. Consider the set function $\mu\colon\mathscr{A}\to[0,1]$, and the system of inequalities in the unknown possibility distribution π on X:*

$$\sup_{x\in A}\pi(x)\leq\mu(A),\quad A\in\mathscr{A}.$$

The following statements hold.

(i) *The system of inequalities has a greatest solution π^g, given by*

$$\pi^g(x)=\inf_{A\in\mathscr{A},x\in A}\mu(A),\quad x\in X.$$

It has a normal solution if and only if π^g is normal.

(ii) *The corresponding system of equalities has a solution if and only if π^g is its greatest solution. It has a normal solution if and only if π^g is normal.*

Proof. To prove the first statement, we first show that π^g is a solution. Consider $A\in\mathscr{A}$ and $x\in A$. Then $\pi^g(x)=\inf_{B\in\mathscr{A},x\in B}\mu(B)\leq\mu(A)$, so $\sup_{x\in A}\pi^g(x)=\sup_{x\in A}\inf_{B\in\mathscr{A},x\in B}\mu(B)\leq\mu(A)$, and therefore π^g is a solution of the system of inequalities. Next, let π be any other solution. Consider x in X. For any $A\in\mathscr{A}$ such that $x\in A$ it follows from the inequalities that $\pi(x)\leq\mu(A)$, so $\pi(x)\leq\inf_{A\in\mathscr{A},x\in A}\mu(A)=\pi^g(x)$. So π^g is the greatest solution of the system of inequalities. Since $\sup\pi^g\geq\sup\pi$ the system has a normal solution if and only if π^g is normal, i.e., if $\sup\pi^g=1$. Since any solution of the corresponding system of equalities is also a solution of the system of inequalities, the proof of the second statement is now trivial. □

Lemma 42. *Let X be an arbitrary gamble on Ω and let $x\in\mathbb{R}$. Then*

$$\mathscr{N}(X-x)\neq\emptyset\Leftrightarrow x\in[\inf X,\sup X].$$

Proof. Consider the evaluation functional X^* on the topological dual space $\mathscr{L}(\Omega)^*$ of continuous linear functionals on $\mathscr{L}(\Omega)$, associated with the gamble X. Obviously, $\mathscr{N}(X-x) \neq \emptyset \Leftrightarrow x \in X^*(\mathbb{P})$. Since $\mathbb{P}$ is convex and X^* is linear, the linear image $X^*(\mathbb{P})$ is a convex subset of $\mathbb{R}$. Since X^* is weak*-continuous and $\mathbb{P}$ is weak*-compact, the continuous image $X^*(\mathbb{P})$ is a compact and therefore closed and bounded subset of $\mathbb{R}$. This tells us that $X^*(\mathbb{P})$ is a closed and bounded real interval. We now show that $\mathscr{N}(X-x)$ is empty if $x < \inf X$. For such x there is a $\delta > 0$ such that $x+\delta < \inf X$. Consequently, $X(\omega)-x \geq \delta$, $\omega \in \Omega$, whence, for any linear prevision P in $\mathbb{P}$, $P(X)-x = P(X-x) \geq P(\delta) = \delta > 0$. Consequently $P(X) > x$. A completely analogous course of reasoning tells us that $\mathscr{N}(X-x)$ is empty if $\sup X < x$. To complete the proof, we show that $\mathscr{N}(X-x)$ is non-empty if $\inf X < x < \sup X$. Indeed, for such x there are ω_1 and ω_2 in Ω such that $X(\omega_1) < x < X(\omega_2)$. Consider the (degenerate) linear prevision P_{ω_1} defined by $P_{\omega_1}(Y) = Y(\omega_1)$, $Y \in \mathscr{L}(\Omega)$. The (degenerate) linear prevision P_{ω_2} is defined similarly. Let $\lambda = [x-X(\omega_1)]/[X(\omega_2)-X(\omega_1)]$. Then $x = (1-\lambda)X(\omega_1)+\lambda X(\omega_2)$ and $\lambda \in (0,1)$. The convex mixture $P = (1-\lambda)P_{\omega_1}+\lambda P_{\omega_2}$ is again a linear prevision, and it satisfies $P(X) = x$, whence $P \in \mathscr{N}(X-x)$. □

APPENDIX B. SOME COMMENTS ON FULLNESS

In this appendix, I have gathered a number of interesting technical results concerning the notion of fullness for arbitrary $\mathbb{P}-[0,1]$-maps, which are essentially candidate second-order possibility distributions. Let us denote the set of these maps by $\mathbb{D}$. Recall that such a map π is full if and only if it is quasi-concave and upper semi-continuous, i.e., if and only if its cut sets $\pi_\alpha = \{P \in \mathbb{P} \colon \pi(P) \geq \alpha\}$ are convex, weak*-closed subsets of $\mathbb{P}$ for all $\alpha \in (0,1]$. I now show that given $\pi \in \mathbb{D}$, there is a smallest full element of $\mathbb{D}$ that dominates π (given the point-wise ordering $\leq$ on $\mathbb{D}$). This element is denoted by $F(\pi)$, and it is called the *full closure* of π.

Proposition 43. *Let $(\pi_j \mid j \in J)$ be an arbitrary family of full elements of $\mathbb{D}$. Then the point-wise infimum $\pi = \inf_{j\in J}\pi_j$ of this family is full.*

Proof. If $J = \emptyset$ then π is identically equal to 1, so all its cut sets are equal to $\mathbb{P}$, a convex and weak*-closed set. Assume, therefore, that J is non-empty. Observe that for any $\alpha \in (0,1]$, $\pi_\alpha = \bigcap_{j\in J}\pi_{j\,\alpha}$. Since an arbitrary intersection of weak*-closed and convex sets is still weak*-closed and convex, this tells us that π is indeed full. □

As a result

$$F(\pi) = \inf\{\pi' \in \mathbb{D} \colon \pi' \text{ is full and } \pi \leq \pi'\},$$

since the right-hand side is full and dominates π. Let us now restrict ourselves to the set $\mathbb{D}_N$ of normal $\mathbb{P}-[0,1]$-maps, precisely the set of all possible (coherent) second-order possibilistic models. We can consider F as an operator on $\mathbb{D}_N$. Note that F is internal in $\mathbb{D}_N$: any normal π is mapped into a normal (modal!) $F(\pi)$.

I now intend to give a number of alternative expressions for $F(\pi)$. The first one is very intuitive, and is based on the observation that a full possibility distribution has convex and weak*-closed cut sets. For every $\alpha \in (0,1]$, consider the convex closure $\overline{\text{co}}(\pi_\alpha)$ of the cut set π_α, that is, the smallest convex and weak*-closed set that includes π_α. We can use the (decreasingly nested) sets $\overline{\text{co}}(\pi_\alpha)$ to construct a new (normal) possibility distribution $\pi^\#$ that dominates π, as follows:

$$\pi^\#(P) = \sup\{\alpha \in (0,1] \colon P \in \overline{\text{co}}(\pi_\alpha)\}.$$

It is important to note that it does *not* necessarily follow that the cut sets $\pi^\#_\alpha$ of $\pi^\#$ are precisely the $\overline{\text{co}}(\pi_\alpha)$. Verify that, since the $\overline{\text{co}}(\pi_\alpha)$ are nested decreasingly, there is the following general relationship between the $\overline{\text{co}}(\pi_\alpha)$ and the $\pi^\#_\alpha$:

$$\pi^\#_\alpha = \bigcap_{0<\gamma<\alpha} \overline{\text{co}}(\pi_\gamma). \tag{38}$$

Interestingly, this construction leads to the full closure $F(\pi)$ of π, as I now show in Proposition 44. The proposition also gives two other expressions, which are closely related to each other. They look daunting, but become much more manageable when π already has some other properties, as is shown in Proposition 47 further on.

Proposition 44. *Let π be a normal possibility distribution on $\mathbb{P}$. Then*

$$F(\pi)\cdot P = \sup\{\alpha \in (0,1] \colon P \in \overline{\mathrm{co}}(\pi_\alpha)\} \tag{39}$$

$$= \inf_{X\in\mathscr{L}(\Omega)} \inf_{y>P(X)} \sup\{\pi(Q) \colon Q(X) \leq y\} \tag{40}$$

$$= \inf_{X\in\mathscr{L}(\Omega)} \inf_{y>P(X)} \sup\{\pi(Q) \colon Q(X) < y\} \tag{41}$$

Proof. It is easy to prove that the alternative expressions (40) and (41) are identical. Denote any of them by $\pi^*(P)$, so it needs to be proven that $F(\pi) = \pi^* = \pi^\#$. As a first step, we prove that for any $\alpha \in (0,1]$, $\pi^*_\alpha = \pi^\#_\alpha$, or in other words, that $\pi^* = \pi^\#$. It can be proven that (see [41, Theorem 3.6.1]):

$$\overline{\mathrm{co}}(\pi_\alpha) = \{P \in \mathbb{P} \colon (\forall X \in \mathscr{L}(\Omega))(P(X) \geq \inf\{Q(X) \colon Q \in \pi_\alpha\})\},$$

so we find for any $P \in \mathbb{P}$, starting with Eq. (38):

$$\begin{aligned}
&P \in \pi^\#_\alpha\\
&\quad\Leftrightarrow (\forall\gamma < \alpha)(P \in \overline{\mathrm{co}}(\pi_\gamma))\\
&\quad\Leftrightarrow (\forall\gamma < \alpha)(\forall X \in \mathscr{L}(\Omega))(P(X) \geq \inf\{Q(X) \colon Q \in \pi_\gamma\})\\
&\quad\Leftrightarrow (\forall\gamma < \alpha)(\forall X \in \mathscr{L}(\Omega))(\forall y > P(X))(y > \inf\{Q(X) \colon Q \in \pi_\gamma\})\\
&\quad\Leftrightarrow (\forall\gamma < \alpha)(\forall X \in \mathscr{L}(\Omega))(\forall y > P(X))(\exists Q \in \mathbb{P})(y > Q(X) \text{ and } Q \in \pi_\gamma)\\
&\quad\Leftrightarrow (\forall X \in \mathscr{L}(\Omega))(\forall y > P(X))(\forall\gamma < \alpha)(\exists Q \in \mathbb{P})(y > Q(X) \text{ and } \pi(Q) \geq \gamma)\\
&\quad\Leftrightarrow (\forall X \in \mathscr{L}(\Omega))(\forall y > P(X))(\forall\gamma < \alpha)(\exists Q \in \mathbb{P})(y > Q(X) \text{ and } \pi(Q) > \gamma)\\
&\quad\Leftrightarrow (\forall X \in \mathscr{L}(\Omega))(\forall y > P(X))(\forall\gamma < \alpha)(\sup\{\pi(Q) \colon Q \in \mathbb{P} \text{ and } y > Q(X)\} > \gamma)\\
&\quad\Leftrightarrow (\forall X \in \mathscr{L}(\Omega))(\forall y > P(X))(\sup\{\pi(Q) \colon Q \in \mathbb{P} \text{ and } y > Q(X)\} \geq \alpha)\\
&\quad\Leftrightarrow P \in \pi^*_\alpha,
\end{aligned}$$

also using Eq. (41). This tells us that $\pi^* = \pi^\#$.

Since each cut set $\pi^\#_\alpha$ is an intersection of convex and weak*-closed sets, by Eq. (38), it is convex and weak*-closed, so $\pi^* = \pi^\#$ is full. Next, we show that $\pi \leq \pi^*$: for any $X \in \mathscr{L}(\Omega)$ and any $y > P(X)$, $\pi(P) \leq \sup\{\pi(Q) \colon y > Q(X)\}$, so indeed

$$\pi(P) \leq \inf_{X\in\mathscr{L}(\Omega)} \inf_{y>P(X)} \sup\{\pi(Q) \colon y > Q(X)\} = \pi^*(P).$$

Since π^* is full, we find that $F(\pi) \leq \pi^*$. This also tells us that starring is an internal operation in $\mathbb{D}_N$. It is furthermore obvious that if we take π_1 and π_2 in $\mathbb{D}_N$, then $\pi_1 \leq \pi_2$ implies $\pi_1^* \leq \pi_2^*$, so starring is an increasing operation. Next, we show that if π is full to start with, then $\pi^* = \pi$. This is easiest if we remember that $\pi^* = \pi^\#$. If π is full, its cut sets π_α are convex and weak*-closed, so $\overline{\mathrm{co}}(\pi_\alpha) = \pi_\alpha$, whence, by Eq. (38), $\pi^\#_\alpha = \bigcap_{0<\gamma<\alpha} \pi_\gamma = \pi_\alpha$. Consequently, $\pi^* = \pi^\# = \pi$.

Finally, it follows from the properties we have proven for starring that, since $\pi \leq F(\pi) \leq \pi^*$, we have $\pi^* \leq F(\pi)^* \leq (\pi^*)^* = \pi^*$, whence $\pi^* = F(\pi)^* = F(\pi)$. □

This allows us to prove an interesting property involving a normal possibility distribution π on $\mathbb{P}$ and its full closure $F(\pi)$.

Proposition 45. *Let π be a normal possibility distribution on $\mathbb{P}$, and let X be a gamble on Ω. Consider the real-valued maps f and g defined on $(0,1]$ by*

$$f(\alpha) = \sup\{P(X) \colon P \in \overline{\mathrm{co}}(\pi_\alpha)\} \quad \textit{and} \quad g(\alpha) = \sup\{P(X) \colon P \in F(\pi)_\alpha\}$$

for all $\alpha \in (0,1]$. *Then* f *and* g *are non-increasing and for all* $\alpha \in (0,1]$, $f(\alpha-) = g(\alpha-)$. *Moreover,* f *and* g *are Riemann-integrable and* $\int_0^1 f(\alpha)\,\mathrm{d}\alpha = \int_0^1 g(\alpha)\,\mathrm{d}\alpha$.

Proof. It is obvious that f and g are non-increasing. They can therefore be discontinuous in at most a countable number of points of their domain $(0,1]$. Therefore, f and g are Riemann-integrable and moreover, $\int_0^1 f(\alpha)\mathrm{d}\alpha = \int_0^1 f(\alpha-)\mathrm{d}\alpha$ and $\int_0^1 g(\alpha)\mathrm{d}\alpha = \int_0^1 g(\alpha-)\mathrm{d}\alpha$. Consider $\alpha \in (0,1]$. If we can show that $f(\alpha-) = g(\alpha-)$ then the proof is complete. Let ε be any real number such that $0 < 2\varepsilon < \alpha$. On the one hand, it follows from $P \in \overline{\mathrm{co}}(\pi_{\alpha-2\varepsilon})$ using Eq. (39) that $F(\pi)\cdot P \geq \alpha - 2\varepsilon$, whence $\overline{\mathrm{co}}(\pi_{\alpha-2\varepsilon}) \subseteq F(\pi)_{\alpha-2\varepsilon}$, and consequently $f(\alpha-2\varepsilon) \leq g(\alpha-2\varepsilon)$. On the other hand, it follows from Proposition 44 and Eq. (38) that if $P \in F(\pi)_{\alpha-\varepsilon}$ then in particular $P \in \overline{\mathrm{co}}(\pi_{\alpha-2\varepsilon})$, whence $F(\pi)_{\alpha-\varepsilon} \subseteq \overline{\mathrm{co}}(\pi_{\alpha-2\varepsilon})$ and consequently $g(\alpha-\varepsilon) \leq f(\alpha-2\varepsilon)$. Combining both inequalities yields $g(\alpha-\varepsilon) \leq f(\alpha-2\varepsilon) \leq g(\alpha-2\varepsilon)$, and if we let $\varepsilon \to 0$, we find that $g(\alpha-) \leq f(\alpha-) \leq g(\alpha-)$, whence indeed $f(\alpha-) = g(\alpha-)$. □

The expressions (40) and (41) for $F(\pi)$ simplify if π is already upper semi-continuous (but not necessarily quasi-concave). To prove this result in Proposition 47 below, we need a general result relating upper semi-continuity and outer regularity of possibility measures. This has been proven in a more general context by Hugo Janssen (private communication). The proof given here is essentially a simplification of his more general proof.

To sketch the context, let X be a non-empty set, and let Π be a possibility measure on X with distribution π. Consider a topology $\mathfrak{T}$ (a collection of open sets) on X, and say that Π is *outer regular* with respect to $\mathfrak{T}$ in a subset A of X if and only if

$$\Pi(A) = \inf\{\Pi(O) \colon A \subseteq O \in \mathfrak{T}\}.$$

Π will be called outer regular if it is outer regular in any subset of X. There is the following interesting link between the outer regularity of Π and the upper semi-continuity of its distribution.

Proposition 46. *Let* Π *be a possibility measure on* X *with distribution* π. *Then the following statements are equivalent:*

(i) Π *is outer regular with respect to* $\mathfrak{T}$;
(ii) Π *is outer regular with respect to* $\mathfrak{T}$ *in the singletons of* X;
(iii) π *is upper semi-continuous with respect to* $\mathfrak{T}$.

Proof. We first prove the equivalence of the first two statements. It clearly suffices to prove that the second implies the first. Assume that Π is outer regular on the singletons of X. Consider an arbitrary subset A of X, then we want to show that $\Pi(A) = \inf\{\Pi(O) \colon A \subseteq O \in \mathfrak{T}\}$. It is obvious from the monotonicity of Π that $\Pi(A) \leq \inf\{\Pi(O) \colon A \subseteq O \in \mathfrak{T}\}$. Assume *ex absurdo* that $\Pi(A) < \inf\{\Pi(O) \colon A \subseteq O \in \mathfrak{T}\} = \lambda$. Then there is an $\varepsilon > 0$ such that $\Pi(A) = \sup_{x\in A}\pi(x) < \lambda - \varepsilon$. It follows from the outer regularity of Π in the singletons that for every $x \in A$, there is an open set $O_x \in \mathfrak{T}$ such that $x \in O_x$ and $\Pi(O_x) < \lambda - \varepsilon$. Let $O = \bigcup_{x\in A} O_x$, then $A \subseteq O$ and $O \in \mathfrak{T}$, so $\Pi(O) \geq \lambda$. On the other hand, $\Pi(O) = \sup_{x\in A}\Pi(O_x) \leq \lambda - \varepsilon$, a contradiction.

Next, we prove the equivalence between the second and third statements. Consider $\alpha \in [0,1)$. We have for any x in X, since Π is outer regular in $\{x\}$:

$$\begin{aligned} x \in \pi_\alpha &\Leftrightarrow \pi(x) \geq \alpha \\ &\Leftrightarrow \inf\{\Pi(O) \colon x \in O \in \mathfrak{T}\} \geq \alpha \\ &\Leftrightarrow (\forall O \in \mathfrak{T})(x \in O \Rightarrow \Pi(O) \geq \alpha) \\ &\Leftrightarrow (\forall O \in \mathfrak{T})(\Pi(O) < \alpha \Rightarrow x \in \mathrm{co}O) \\ &\Leftrightarrow x \in \bigcap_{O\in\mathfrak{T},\Pi(O)<\alpha} \mathrm{co}O \end{aligned}$$

which tells us that π_α is an intersection of closed sets and therefore closed. This implies that π is upper semi-continuous with respect to $\mathfrak{T}$. Conversely, assume that π is upper semi-continuous. This is equivalent with, for any x in X:

$$(\forall\varepsilon > 0)(\exists x \in O \in \mathfrak{T})(\forall y \in O)(\pi(y) < \pi(x)+\varepsilon)$$

whence

$$(\forall\varepsilon > 0)(\exists x \in O \in \mathfrak{T})(\Pi(O) \leq \pi(x)+\varepsilon)$$

which implies that

$$(\forall\varepsilon > 0)(\inf\{\Pi(O)\colon x \in O \in \mathfrak{T})\} \leq \pi(x)+\varepsilon)$$

and therefore also

$$\inf\{\Pi(O)\colon x \in O \in \mathfrak{T})\} \leq \pi(x).$$

Since the converse inequality holds by the monotonicity of Π, it follows that Π is outer regular in $\{x\}$. □

Proposition 47. *Let the normal possibility distribution π on $\mathbb{P}$ be upper semi-continuous. Then for all X in $\mathscr{L}(\Omega)$ and x in $\mathbb{R}$:*

$$\inf_{y>x}\sup\{\pi(Q)\colon Q(X) < y\} = \inf_{y>x}\sup\{\pi(Q)\colon Q(X) \leq y\} = \sup\{\pi(Q)\colon Q(X) \leq x\},$$

and consequently, for any $P \in \mathbb{P}$,

$$F(\pi)\cdot P = \inf_{X\in\mathscr{L}(\Omega)}\sup\{\pi(Q)\colon Q(X) \leq P(X)\}.$$

Proof. Since it always holds that

$$\inf_{y>x}\sup\{\pi(Q)\colon Q(X) < y\} = \inf_{y>x}\sup\{\pi(Q)\colon Q(X) \leq y\} \geq \sup\{\pi(Q)\colon Q(X) \leq x\},$$

it suffices to prove that under the given assumptions,

$$\sup\{\pi(Q)\colon Q(X) \leq x\} \geq \inf_{y>x}\sup\{\pi(Q)\colon Q(X) < y\}.$$

Since π is upper semi-continuous, Proposition 46 tells us that Π is outer regular with respect to the relativisation to $\mathbb{P}$ of the weak* topology on $\mathscr{L}(\Omega)^*$. Call this relativisation $\mathfrak{T}^*$, for ease of notation. Note that $\mathbb{P}$ is weak*-compact, and therefore also compact in this relativisation. Then it follows that

$$\sup\{\pi(Q)\colon Q(X) \leq x\} = \inf\{\Pi(O)\colon \{Q \in \mathbb{P}\colon Q(X) \leq x\} \subseteq O \in \mathfrak{T}^*\}.$$

It only remains to prove that for any open set $O \in \mathfrak{T}^*$ such that $\{Q \in \mathbb{P}\colon Q(X) \leq x\} \subseteq O$ there is a $y > x$ such that $\{Q \in \mathbb{P}\colon Q(X) < y\} \subseteq O$. *Ex absurdo*, assume that for all $y > x$, $\{Q \in \mathbb{P}\colon Q(X) < y\} \cap \mathrm{co}O \neq \emptyset$, so there is a $Q_y \in \mathrm{co}O$ such that $x < Q_y(X) < y$. In particular, this means that there is a sequence of $(Q_n)_{n\in\mathbb{N}}$ of elements of co$O$, such that

$$(\forall n \in \mathbb{N})(x < Q_n(X) < x+\frac{1}{n}). \tag{42}$$

Since O is the intersection of a weak*-open subset O' of $\mathscr{L}(\Omega)^*$ with $\mathbb{P}$, $\mathrm{co}O = \mathbb{P}\setminus O = \mathscr{L}(\Omega)^*\setminus O'$ is a weak*-closed subset of the weak*-compact set $\mathbb{P}$, and is therefore itself weak*-compact. So $(Q_n)_{n\in\mathbb{N}}$ has a weakly* convergent subsequence $(Q_{k(n)})_{n\in\mathbb{N}}$, where k is some order-preserving injection of $\mathbb{N}$ into $\mathbb{N}$ (note that $k(n) \geq n$). We denote the limit of this subsequence by R, and of course $R \in \mathrm{co}O$, since co$O$ is weak*-closed. Since $Q_{k(n)}$ weakly* converges to R, the real sequence $Q_{k(n)}(X)$ converges to $R(X)$, which implies that $R(X) = x$, taking into account Eq. (42). Consequently, $R \in \{Q \in \mathbb{P}\colon Q(X) \leq x\} \subseteq O$, which contradicts $R \in \mathrm{co}O$. □

A special case of this result obtains when π has finite support, i.e. when $\pi(P)$ is zero on all but a finite number $\{P_1, P_2, \dots, P_n\}$ of elements P of $\mathbb{P}$. Verify that

$$F(\pi)\cdot P = \inf_{X\in\mathscr{L}(\Omega)} \max\{\pi(P_k)\colon P_k(X) \leq P(X), k = 1,\dots,n\},$$

and consequently, for $\alpha \in (0,1]$,

$$F(\pi)\cdot P \geq \alpha \Leftrightarrow (\forall X \in \mathscr{L}(\Omega))(P(X) \geq \underline{P}_\alpha(X)),$$

i.e., $F(\pi)_\alpha = \mathscr{M}(\underline{P}_\alpha)$, where $\underline{P}_\alpha(X) = \min\{P_k(X)\colon \pi(P_k) \geq \alpha, k = 1,\dots,n\}$.

REFERENCES

[1] J. O. Berger. The robust Bayesian viewpoint. In J. B. Kadane, editor, *Robustness of Bayesian Analyses*. Elsevier Science, Amsterdam, 1984.

[2] J. O. Berger. *Statistical Decision Theory and Bayesian Analysis*. Springer-Verlag, New York, 1985.

[3] K. P. S. Bhaskara Rao and M. Bhaskara Rao. *Theory of Charges*. Academic Press, London, 1983.

[4] L. Boyen, G. de Cooman, and E. E. Kerre. On the extension of P-consistent mappings. In G. de Cooman, D. Ruan, and E. E. Kerre, editors, *Foundations and Applications of Possibility Theory – Proceedings of FAPT '95*, pages 88–98, Singapore, 1995. World Scientific.

[5] G. de Cooman. Possibility theory I: the measure- and integral-theoretic groundwork. *International Journal of General Systems*, 25:291–323, 1997.

[6] G. de Cooman. Possibility theory II: conditional possibility. *International Journal of General Systems*, 25:325–351, 1997.

[7] G. de Cooman. Possibility theory III: possibilistic independence. *International Journal of General Systems*, 25:353–371, 1997.

[8] G. de Cooman. Possibilistic previsions. In *Proceedings of IPMU '98*, volume I, pages 2–9. Éditions EDK, Paris, 1998.

[9] G. de Cooman. Lower desirability functions: a convenient imprecise hierarchical uncertainty model. In G. de Cooman, F. G. Cozman, S. Moral, and P. Walley, editors, *ISIPTA '99 – Proceedings of the First International Symposium on Imprecise Probabilities and Their Applications*, pages 111–120. Imprecise Probabilities Project, Ghent, 1999.

[10] G. de Cooman. Belief models: an order-theoretic analysis. In G. de Cooman, T. L. Fine, and T. Seidenfeld, editors, *ISIPTA '01 – Proceedings of the Second International Symposium on Imprecise Probabilities and Their Applications*, pages 93–103. Shaker Publishing, Maastricht, 2000.

[11] G. de Cooman. Integration in possibility theory. In M. Grabisch, T. Murofushi, and M. Sugeno, editors, *Fuzzy Measures and Integrals – Theory and Applications*, pages 124–160. Physica-Verlag (Springer), Heidelberg, 2000.

[12] G. de Cooman. Integration and conditioning in numerical possibility theory. *Annals of Mathematics and Artificial Intelligence*, 32:87–123, 2001.

[13] G. de Cooman. Precision–imprecision equivalence in a broad class of imprecise hierarchical uncertainty models. *Journal of Statistical Planning and Inference*, 105:175–198, 2002.

[14] G. de Cooman. Belief models: an order-theoretic investigation. *Journal of Mathematical Analysis and Artificial Intelligence*, 2003. Accepted for publication.

[15] G. de Cooman and D. Aeyels. Supremum preserving upper probabilities. *Information Sciences*, 118:173–212, 1999.

[16] G. de Cooman and D. Aeyels. A random set description of a possibility measure and its natural extension. *IEEE Transactions on Systems, Man and Cybernetics—Part A: Systems and Humans*, 30:124–130, 2000.

[17] G. de Cooman and P. Walley. A possibilistic hierarchical model for behaviour under uncertainty. *Theory and Decision*, 52:327–374, 2002.

[18] B. de Finetti. La prévision: ses lois logiques, ses sources subjectives. *Annales de l'Institut Henri Poincaré*, 7:1–68, 1937. English translation in [37].

[19] B. de Finetti. *Teoria delle Probabilità*. Einaudi, Turin, 1970.

[20] B. de Finetti. *Theory of Probability*, volume 1. John Wiley & Sons, Chichester, 1974. English Translation of [19].

[21] B. de Finetti. *Theory of Probability*, volume 2. John Wiley & Sons, Chichester, 1975. English Translation of [19].

[22] D. Dubois, J. Lang, and H. Prade. Possibilistic logic. In D. M. Gabbay, C. J. Hogger, and J. A. Robinson, editors, *Handbook of Logic in Artificial Intelligence and Logic Programming*, volume 3, pages 439–513. Oxford University Press, 1994.

[23] D. Dubois and H. Prade. Additions of interactive fuzzy numbers. *IEEE Transactions on Automatic Control*, 26:926–936, 1981.

[24] D. Dubois and H. Prade. *Théorie des possibilités*. Masson, Paris, 1985.

[25] D. Dubois and H. Prade. *Possibility Theory*. Plenum Press, New York, 1988.

[26] P. Gärdenfors and N.-E. Sahlin. Unreliable probabilities, risk taking, and decision making. *Synthese*, 53:361–386, 1982.
[27] P. Gärdenfors and N.-E. Sahlin. *Decision, Probability, and Utility*. Cambridge University Press, Cambridge, 1988.
[28] L. Gilbert, G. de Cooman, and E. E. Kerre. Practical implementation of possibilistic probability mass functions. *Soft Computing*, 7:304–309, 2003.
[29] M. Goldstein. The prevision of a prevision. *Journal of the American Statistical Society*, 87:817–819, 1983.
[30] I. J. Good. Subjective probability as the measure of a non-measurable set. In E. Nagel, P. Suppes, and A. Tarski, editors, *Logic, Methodology and Philosophy of Science*, pages 319–329. Stanford University Press, Stanford, 1962.
[31] I. J. Good. Some history of the hierarchical Bayesian methodology. In J. M. Bernardo, M. H. DeGroot, D. V. Lindley, and A. F. M. Smith, editors, *Bayesian Statistics*, volume 1, pages 489–519. Valencia University Press, Valencia, 1980.
[32] J. Halliwell and Q. Shen. Towards a linguistic probability theory. In *Proceedings of the 11th International Conference on Fuzzy Systems*, pages 596–601. 2002.
[33] R. B. Holmes. *Geometric Functional Analysis and Its Applications*. Springer-Verlag, New York, 1975.
[34] P. J. Huber. *Robust Statistics*. John Wiley & Sons, New York, 1981.
[35] E. E. Kerre. Basic principles of fuzzy set theory for the representation and manipulation of imprecision and uncertainty. In E. E. Kerre, editor, *Introduction to the Basic Principles of Fuzzy Set Theory and Some of Its Applications*, pages 1–158. Communication & Cognition, Ghent, 1991.
[36] H. E. Kyburg. Higher order probabilities and intervals. *International Journal of Approximate Reasoning*, 2:195–209, 1988.
[37] H. E. Kyburg Jr. and H. E. Smokler, editors. *Studies in Subjective Probability*. Wiley, New York, 1964. Second edition (with new material) 1980.
[38] C. V. Negoita and D. A. Ralescu. *Applications of Fuzzy Sets to Systems Analysis*. Birkhauser Verlag, Basel, 1975.
[39] F. P. Ramsey. Truth and probability. In R. B. Braithwaite, editor, *The Foundations of Mathematics*, pages 156–198. Routledge & Kegan Paul, London, 1931. Reprinted in [37] and [27].
[40] G. Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, Princeton, NJ, 1976.
[41] P. Walley. *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, London, 1991.
[42] P. Walley. Statistical inferences based on a second-order possibility distribution. *International Journal of General Systems*, 26:337–383, 1997.
[43] P. Walley and G. de Cooman. A behavioural model for linguistic uncertainty. *Information Sciences*, 134:1–37, 2001.
[44] Z. Wang and G. J. Klir. *Fuzzy Measure Theory*. Plenum Press, New York, 1992.
[45] P. M. Williams. Notes on conditional previsions. Technical report, School of Mathematical and Physical Science, University of Sussex, UK, 1975.
[46] L. A. Zadeh. Fuzzy sets. *Information and Control*, 8:338–353, 1965.
[47] L. A. Zadeh. The concept of a linguistic variable and its application to approximate reasoning I. *Information Sciences*, 8:199–249, 1975.
[48] L. A. Zadeh. The concept of a linguistic variable and its application to approximate reasoning II. *Information Sciences*, 8:301–357, 1975.
[49] L. A. Zadeh. The concept of a linguistic variable and its application to approximate reasoning III. *Information Sciences*, 9:43–80, 1976.
[50] L. A. Zadeh. Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, 1:3–28, 1978.
[51] L. A. Zadeh. Fuzzy probabilities. *Information Processing and Management*, 20:363–372, 1984.
[52] L. A. Zadeh. Toward a perception-based theory of probabilistic reasoning with imprecise probabilities. *Journal of Statistical Planning and Inference*, 105:233–264, 2002.

GHENT UNIVERSITY, SYSTEMS RESEARCH GROUP, TECHNOLOGIEPARK – ZWIJNAARDE 914, 9052 ZWIJNAARDE, BELGIUM
E-mail address: gert.decooman@ugent.be