Confidence and ambiguity*

Brian Hill
GREGHEC, HEC Paris†
February 20, 2009

Abstract
This paper proposes a model of the decision-maker’s confidence in his probability judgements, in terms of an implausibility measure – a real-valued function on the set of probability functions. A decision rule is axiomatised according to which the decision-maker evaluates acts using sets of probability functions which vary depending on the agent’s implausibility measure and on what is at stake in the choice of the act. The framework proposed yields a natural notion of comparative aversion to lack of confidence, or ambiguity aversion, and allows the definition of an ambiguity premium. It is shown that these notions are equivalent and can be characterised in terms of the implausibility measure representing the agent’s confidence. A simple portfolio example is presented.

Keywords: Confidence; multiple priors; ambiguity aversion; ambiguity premium; implausibility measure.

JEL classification: D81, C69.

1 Introduction
Ever since the seminal paper of Ellsberg (1961), and with increased interest in recent years, the problem of choice under “ambiguity” has occupied decision theorists. Without wishing to enter into the debate about the correct definition of ambiguity, and the correct model of choice under ambiguity (Epstein and Zhang, 2001; Ghirardato et al., 2004; Wakker, 2001), we shall take a decision under ambiguity to be a case where there might be uncertainty regarding the correct probabilities to use in the decision. As such, the problem of choice under ambiguity seems to bring in the question of the decision-maker’s confidence about his probability judgements. He is fully confident that the probability that a fair die falls on a one is a sixth: choice under risk is a case of full confidence. By contrast, although he thinks that the odds that the shares of a

---

*We would like to thank Itzhak Gilboa and Edi Karni for stimulating discussion and helpful comments.
†GREGHEC, HEC Paris. 1 rue de la Libération, 78351 Jouy-en-Josas, France. Tel: +33 1 39 67 72 65. Fax: +33 1 39 67 71 09. E-mail: hill@hec.fr
particular company B will rise are better than 0.8, he is not entirely confident in this judgement (he acknowledges that it may be that the odds are only around 0.5): this is a case of uncertainty about the probabilities, i.e. of choice under ambiguity.

In the light of this relationship, the problem of understanding choice under ambiguity can be resumed in the following two questions: What is the best way to represent the agent’s confidence in his probability judgements? What role does his confidence play in his decisions? A popular family of models, initiated by Gilboa and Schmeidler (1989), use a set of probability functions in response to the first question. Several rules have been proposed which specify the role of such sets of probabilities in decision-making (Gilboa and Schmeidler, 1989; Bewley, 2002; Ghirardato et al., 2004). In these models, confidence seems to be an all or nothing affair: the decision-maker is fully confident of any probability judgement which is true according to all of the probability functions in this set, and does not forward any probability judgements of which he is not fully confident. Gajdos et al. (2008) develop this representation by supposing that the appropriate set of probability functions is itself a function of an exogenously given set of probability functions, to be thought of as the “objective” information the decision-maker has received. In their model, confidence is still an all or nothing affair, but rather than being fixed, the probability judgements in which the decision-maker is (fully) confident are determined by his information.

In this paper, we propose a model according to which confidence comes in degrees: the decision-maker is more or less confident of different probability judgements. Using such a model, we propose a decision rule according to which the evaluation of a prospect does not use a fixed set of probability functions (as in Gilboa and Schmeidler (1989)), nor a set which is determined by some exogenously given factor (as in Gajdos et al. (2008)), but a set which varies with the confidence appropriate for the prospect under consideration. The guiding intuition behind the rule is that prospects in which the “stakes” are larger require more confidence.

Consider firstly the representation of the decision-maker’s confidence. The basic idea stems from the following truism: the more confident one is in a probability judgement, the more tenaciously one subscribes to it or, equivalently, the less willing one is to forgo it. If the agent is confident that the probability that the shares of a particular company A will rise is greater than 0.7, he will use this aspect extensively in your decisions, where appropriate; by contrast, if he thinks that the probability that the shares of company B will rise is greater than 0.8, but he is not very confident of his estimate, he is less willing to rely on this belief in important decisions. Construed thus, confidence can be modelled by what we shall call a implausibility measure on the space of probability functions, which assigns a real number (or infinity) to each probability function, representing its degree of implausibility. One probability function is at the centre, i.e. has implausibility 0: this is the agent’s best estimate of the probabilities of the relevant events. The other probability functions are closer or farther away from the centre: those which are closer are serious candidates for the correct probability function, those which are farther are more implausible. The agent’s confidence in a probability judgement is given by this implausibility measure: the farther he has to go to find a probability function which disagrees with the judgement, the more confidence he has in it. If there are “close” probability functions which contradict the probability judgement, his confidence in the judgement is low.
The idea is represented graphically in Figure 1:¹ the points are probability functions, and the concentric spheres contain probability functions which have an implausibility less than a certain value. The larger confidence in the probability judgement concerning company A is represented by the fact that the largest sphere containing only probability functions confirming this judgement (the light grey sphere) contains the largest sphere containing probability functions which all confirm the judgement concerning company B. This corresponds to the fact that the judgement concerning A is confirmed by all probabilities up to a higher level of implausibility ($d_1$) that the judgement concerning B (where the implausibility level is $d_2$). Naturally, the implausibility measure is a subjective element, describing the agent’s attitude (see Section 4.3.2).

It remains to specify how such a representation of the agent’s confidence in his probability judgements is involved in his choices. The intuition is simple. As we said, the less confident an agent is in a probability judgement, the less willing he is to hang on to and employ that judgement; accordingly, the less confident one is in a probability judgement, the lower the stakes at which one is willing to base one’s choice (or bet) on this judgement. If the agent is very confident that the probability that company A’s shares rise is greater than 0.7, he will be willing to bet the monetary equivalent of $u$ units of utility on the share-price going up, even for large values of $u$.² On the other hand, if he thinks that the probability of company B’s shares rising is greater than 0.8, but he is not very confident in this judgement, one might expect different behaviour: whilst he may be willing to bet the monetary equivalent of $u$ units of utility on the price increase for B.

---

¹Where $p(A)$ is short for “the probability that the shares of company A rise” and similarly for B.
²By bet, we mean the simple prospect where the agent gains $u$ units of utility if the shares go up and loses $u$ units if they do not; preferring to bet means that this prospect is preferred to a sure zero change in utility.
going up for small $u$, he will be reticent to place such bets for larger $u$. He might be willing to bet a beer on the movement in the stock price, but he wouldn’t bet his fortune. The higher the stakes the decision-maker is willing to accept, the more confident he is in the probability judgement underlying his choice; this intuition is behind the main representation result in this paper (Section 2.2). Note that, to measure the stakes, we do not use the (usually monetary) outcomes involved in the bet, but the agent’s utility values for these outcomes; this ensures that it is the agent’s attitudes to his confidence, or to ambiguity, which are being considered, not his attitudes to risk.

Cashing this out, if the stakes involved in the choice of an option $f$ are $\sigma(f)$, then the agent will use all the probability functions with degree of implausibility less than $\sigma(f)$ in his evaluation of $f$. We assume that he uses the maxmin decision rule over this set of probability functions, so that he (weakly) prefers option $g$ to option $f$ if and only if:

\[
\lvert(1)\rvert\quad \min_{p \in S \text{ s.t. } d(p) \leq \sigma(f)} \sum_{s \in S} u(f(s))p(s) \leq \min_{p \in S \text{ s.t. } d(p) \leq \sigma(g)} \sum_{s \in S} u(g(s))p(s)
\]

The abstract example considered above is suggestive of many applications; we shall content ourselves with mentionning a few. For one, the probabilities and confidence in the example above are typical representations of an investor’s attitudes towards a company A which is in his home market and company B which is quoted on a foreign market. Although he might think the home stock to be less likely to be lucrative than the foreign stock, he is more confident in his judgement regarding the former; as indicated above, under the proposed decision rule, this may translate to a higher investment in the home stock than in the foreign one. Of course, this sort of behaviour has been observed: it is the so-called “home-bias”. We shall analyse a simple example of this sort in Section 5. Another potential application concerns expert advice. People often make decisions, and give advice, if the stakes are low (a beer, in the example, or somebody else’s money), which they would not make, or follow, if the stakes were higher (the money was their own). The model proposed here can account for this sort of behaviour: though they are giving their best estimates, they are not confident enough in the probability judgements underlying their advice to act upon it themselves.

In Section 2 we introduce the framework and basic definitions, and we prove the main theorem, which gives sufficient conditions for the existence of a (suitably) unique implausibility measure and utility function such that the agent’s preferences are represented by (1). In Section 3 we define a notion of comparative aversion to lack of confidence, which can be thought of as a sort of notion of comparative ambiguity aversion. We compare it to existing notions of comparative ambiguity aversion, define an ambiguity premium, and offer a characterisation of these notions, in the style of an the Arrow-Pratt characterisation of risk aversion. Section 4 is dedicated to extensions, variants, comments and discussion of related literature. Section 5 contains an application to a simple portfolio problem.

---

3 $\lfloor x \rfloor = x$ if $x \geq 0$ and 0 otherwise.

4 Thanks to Itzhak Gilboa for this example.
2 Preliminaries and Representation

2.1 Preliminaries

Let $S$ be a non-empty finite set of states, with $\Sigma$ the $\sigma$-algebra of all subsets of $S$, which are called events. $\Delta(\Sigma)$ is the set of finitely-additive probability measures on $(S, \Sigma)$. Where necessary, we shall employ the weak$^*$ topology on the set of finitely-additive set functions; under this topology, a net of $p_\alpha \in \Delta(\Sigma)$ converges to $p$ if and only if $p_\alpha(A) \rightarrow p(A)$ for all $A \in \Sigma$. $X$ is a nonempty set of outcomes; a consequence is a probability measure on $X$ with finite support. $\Delta(X)$ is the set of consequences. (Simple) acts are simple measurable functions from states to consequences; $A$ is the set of simple acts. So, for an act $f$, and a state $s$, $f(s)$ is a lottery over $X$ with finite support; for a utility function $u$ over $X$, we will denote the expected utility of this lottery by $u(f(s)) = \sum_{x \in \text{supp}(f(s))} f(s)(x) u(x)$.

$A$ is a mixture set with the mixture relation defined pointwise: for $f, h$ in $A$, and $a \in \mathbb{R}$, $0 < a < 1$, the mixture $af + (1 - a)h$ is defined by $(af + (1 - a)h)(s, x) = af(s, x) + (1 - a)h(s, x)$ (Fishburn, 1970, Ch 13).\(^5\) With slight abuse of notation, a constant act taking consequence $c$ for every state will be denoted $c$ and the set of constant acts will be denoted $\Delta(X)$. Similarly, both the degenerate lottery taking value $x \in X$ and the constant act yielding this lottery will be denoted $x$, and both the set of degenerate lotteries and the set of constant acts yielding degenerate lotteries will be denoted $X$.

Here we think of the outcomes as monetary values, and thus generally assume that $X$ is the set of real numbers; in particular, it comes equipped with a dense, unbounded weak order $\leq$. We assume a preference relation $\preceq$ on $A$; $\sim$ and $\prec$ are defined to be the symmetric and asymmetric components of $\preceq$. Null events and null states are defined in the usual way.\(^6\)

The following notion, mentioned in the Introduction, will be central.

Definition 1. An implausibility measure on $\Delta(\Sigma)$ is a function $d : \Delta(\Sigma) \rightarrow \mathbb{R}_{\geq 0} \cup \{\infty\}$ which is continuous on $[0, \sup_{p \text{ st. } d(p) \neq \infty} d(p))$ and is such that $d^{-1}(0)$ contains one element. This element is called the centre.

The terminology used here refers to the intuition that the probabilities are closer or further from “acceptability”. It does not refer to any measure-theoretic structure; the word “measure” is only adopted because “degree of implausibility function” is clumsy. Implausibility measures are continuous wherever they take non-maximal non-infinite values. The implausibility measures used here will be centred: there is just one probability function with implausibility zero. This translates the assumption that if the agent was forced to give his best estimate for the probability of any event, he could come up with a single value (and the values he comes up with satisfy the probability axioms). See Section 4 for a discussion of weakenings of this assumption.

The ordinal equivalent of the cardinal notion of implausibility measure is as follows.

\(^5\)Here, Anscombe and Aumann’s “Reversal of Order” axiom is tacitly assumed to hold (Anscombe and Aumann, 1963). As is standard in much work with this framework, this axiom is assumed to hold throughout this paper.

\(^6\)An event $A$ is null if, for any acts $f$ and $f'$ differing only on $A$, $f \sim f'$. A state $s$ is null if the singleton event $\{s\}$ is null.
**Definition 2.** For implausibility measure \( d \) on \( \Delta(\Sigma) \), the order restriction of \( d \), \( \leq^d \) is the order on \( \Delta(\Sigma) \) defined as follows: \( p \leq^d q \) iff \( d(p) \leq d(q) \).

Finally, for \( x \in \mathbb{R} \), we define \( \lfloor x \rfloor \) to be \( x \) if \( x \geq 0 \) and 0 otherwise.

### 2.2 The Representation

We will pose several general axioms, most of which only involve standard properties of preference orders. The standard properties are recalled in Appendix A. First of all:

**Axiom A1 (Basic axioms).** \( \preceq \) satisfies non triviality, weak order, continuity, unboundedness of outcomes – for any \( f \in A \), there exists \( x, y \in X \), \( x \preceq f \preceq y \) – and monotonicity on outcomes – for any \( x, y \in X \), \( x \preceq y \) iff \( x \preceq y \).

By the standard argument, weak order, continuity and unboundedness of outcomes ensure that every act has a certainty equivalent in \( \Delta(X) \) according to \( \preceq \). Monotonicity on outcomes ensures moreover that every act has a certainty equivalent in \( X \): that is, for any act \( f \), there is a unique constant act in \( X \) which is \( \preceq \)-equivalent to it. We will denote it by \( [f] \) and henceforth refer to it as the certainty equivalent of \( f \). The full strength of the monotonicity on outcomes conditions is not necessary, but greatly simplifies the exposition; see Section 4.

**Axiom A2 (Determinate utilities).** The restriction of \( \preceq \) to \( \Delta(X) \) (the set of constant acts) satisfies independence.

By the von Neumann-Morgernstern theorem, this axiom guarantees that there is an expected utility representation of the restriction of \( \preceq \) to constant acts. Let \( u \) be a utility function representing these preferences.

Let us introduce the following definitions.

**Definition 3.** For any \( f \in A \) and any \( r \in \mathbb{R} \), let \( f + r \) be the act taking values in \( X \) such that \( u(f(s) + r) = u(f(s)) + r \) for all \( s \in S \).

\( f + r \) is well-defined because the utility is strictly increasing, continuous and unbounded (by A1).

**Definition 4.** For \( f \in A \), let \( \sigma(f) = -\min_{s \in S \text{ non-null}} u(f(s)) \).

\( \sigma \) is to be interpreted as the “stakes” implied in the choice of \( f \). Roughly, it is the most the agent could lose if he choses \( f \). For further discussion of \( \sigma \), see in particular Sections 4.1.5 and 4.2.2. Note that \( \sigma \) depends on the utility function \( u \).

The basic idea is that the set of probabilities used in the maxmin evaluation of an act will depend on the stakes of the act. Hence, acts with the same stakes will have the same set of probabilities involved in their evaluation: the ordinary maxmin axioms apply on the preferences between acts with the same stakes. However, some of these axioms (for example uncertainty aversion) involve mixtures of acts, and a mixture of two acts with stakes \( r \) may not have stakes \( r \). It is not sufficient to pose the axioms on such a restricted set. Instead, we ask that the maxmin axioms hold on acts, when we evaluate them as if they had the same stakes, say \( r \). That is: to compare any two acts,
first we shift them so that each has stakes $r$; then we “evaluate” each, which, given that we only have preferences, means taking the certainty equivalent; finally, we shift the certainty equivalents back (i.e. undo the shift that brought the stakes of the acts to $r$), and compare the results. The gives the preference between $f$ and $g$ as if they both had stakes $r$. Hence the following definition.

**Definition 5.** For any real number $r$, the order $\preceq_r$ is defined as follows: for any $f$, $g$ in $\mathcal{A}$, $f \preceq_r g$ iff $\left[ f + (\sigma(f) - r) \right] + (r - \sigma(f)) \preceq \left[ g + (\sigma(g) - r) \right] + (r - \sigma(g))$.

For each act $f$, $f + (\sigma(f) - r)$ is the “transform” of $f$ which has stakes $r$, $[f + (\sigma(f) - r)]$ is the certainty equivalent of this transform and $[f + (\sigma(f) - r)] + (r - \sigma(f))$ is the certainty equivalent obtained by “undoing” the shift to stakes $r$.

Note that this definition and the intuition given above make sense because the maxmin representation (which is obtained imposing the axioms below) is a linear functional, so that these shifts do not change the preference order between the acts; i.e. for any $f$, $g$ in $\mathcal{A}$ and any real number $a$, $MMEU_C(f) \leq MMEU_C(g)$ iff $MMEU_C(\quad (f + a) - a \leq MMEU_C(\quad (g + a) - a$ (writing $MMEU_C(f)$ for the largest minimal expected utility of $f$ over a set of probabilities $C$).

**Axiom A3** (MMEU). For all $r$, $\preceq_r$ satisfies monotonicity, $C$-independence and uncertainty aversion.

These are the Gilboa and Schmeidler (1989) maxmin axioms. They imply that $\preceq_r$ can be represented by maximisation of minimum expected utility over a set of probabilities.

We define the “unambiguous preferences” at $r$, for any $r$, as in Ghirardato et al. (2004, Definition 3):

**Definition 6.** For $r$ a real number, the unambiguous preference order at $r$, $\preceq^*_r$ is defined as follows: $f \preceq^*_r g$ if $\lambda f + (1 - \lambda)h \preceq_r \lambda g + (1 - \lambda)h$ for all $\lambda \in (0, 1]$ and all $h$ in $\mathcal{A}$.

Note that this order is not necessarily complete.

This allows us to formulate the axiom which “glues together” the representations obtained using the previous axioms.

**Axiom A4** (Consistency). For all $r, r' \in \mathbb{R}$, if there exists $s \in \mathbb{R}$ with $s \geq r$ and $\preceq^*_r \preceq^*_s \preceq^*_r r' \leq s$ or $s \geq r'$ and $\preceq^*_r \preceq^*_r s = \preceq^*_r r'$, then $r \leq r'$ iff $\preceq^*_r \preceq^*_r r' \leq \preceq^*_r s$.

The following, largely technical, axiom ensures appropriate continuity.

**Axiom A5** (Continuity). For any act $g \in \mathcal{A}$ and any constant act $c \in \Delta(X)$, the sets $\{ r | g \succeq_r c \}$ and $\{ r | g \preceq_r c \}$ are closed.

One final axiom is needed to ensure centering:

**Axiom A6** (Centering). There exists a maximal $\Lambda \in \mathbb{R}$ such that, for any $r, r' \leq \Lambda$, $\preceq_r = \preceq_{r'}$ and both satisfy, in addition to the above axioms, independence.

We will say that a utility function is zeroed if $\Lambda = 0$. Note that, since $\sigma$ and hence the $\preceq_r$ depend on $u$, this is indeed a condition on $u$.

These axioms suffice for a representation of preferences by a decision rule featuring confidence, represented by an implausibility measure.
Theorem 1. If A1-A6, there exists a zeroed utility function $u : X \to \mathbb{R}$ and an implausibility measure on the space of probabilities $d : \Delta(\Sigma) \to \mathbb{R}_{\geq 0} \cup \{\infty\}$ such that, for all $f, g \in A$, $f \succeq g$ iff

$$
\min_{d(p) \leq \sigma(f)} \sum_{s \in S} u(f(s))p(s) \leq \min_{d(p) \leq \sigma(g)} \sum_{s \in S} u(g(s))p(s)
$$

Furthermore, for any other zeroed utility function $u'$ and implausibility measure $d'$ representing $\succeq$ according to (1), there exists a positive real number $a$ such that $u' = au$ and $d' = \frac{1}{a}d$.

Of course, expected utility is the special case where $d(q) = \infty$ if $d(q) = 0$.

3 Measures of confidence and ambiguity aversion

The implausibility measure on the space of probability functions involved in (1) represents the agent’s relative confidence in one probability judgement compared to another: he is more confident if he has to go to a higher implausibility (further away on the diagram in Figure 1) to find a probability function which contradicts his probability judgement. In this capacity, numerical values of distance are not required: the order restriction of the implausibility measure (Definition 2) contains all the necessary information. What the numerical values do provide, as Theorem 1 makes clear, is the link between confidence, utility (and in particular the “stakes” involved) and choice.

There may thus be two agents with the same order on the space of probability functions (and in particular the same centre), the same utility function, but making different choices, because of differing implausibility measures. Consider a pair of acts $f, g$ with the same “stakes” such that $g$ is (weakly) preferred to $f$ if one chooses using expected utility and the centre probability. Suppose that the least implausible rung of the order restriction where the maxmin rule yields a (strict) preference for $f$ over $g$ is the smallest one containing a probability function $q$. The difference between the agents comes out in the “stakes” up to which they are willing to continue to choose (transforms of) $g$ over (transforms of) $f$. The implausibility measures each associates to $q$ are different: the agent who gives $q$ a larger implausibility will choose (transforms of) $g$ up to higher stakes that the one giving it a lower implausibility value. But the stakes one is willing to bet on seem to be a good measure of one’s aversion to one’s lack of confidence in one’s probability judgements. The former agent can go to higher stakes than the latter without his confidence concerning whether probability $q$ should be taken into account entering into play; he is less averse to his lack of confidence. This reasoning provides the following notion of comparative aversion to lack of confidence.

Definition 7. Let $\succeq^1$ and $\succeq^2$ be represented by the same utility and implausibility measures which have the same order restriction on probabilities. Then $\succeq^1$ is more averse to lack of confidence than $\succeq^2$ if, for any $f, g \in A$ and any real number $\rho$, if $f \succeq^1_r g$ for all $r \leq \rho$, then $f \succeq^2_r g$ for any $r \leq \rho$. 

8
To the extent that one’s confidence in one’s probability judgements is related to ambiguity (in the sense specified at the beginning of the paper), this can be thought of as a notion of comparative ambiguity aversion. However, at first sight, it differs from notions of comparative ambiguity aversion which have been proposed in the literature. Here is a definition of comparative ambiguity aversion which mimics closely, in the current framework, some recently proposed definitions.\footnote{This definition is closest to that in Klibanoff et al. (2005), who suppose the same second-order probabilities for the two decision-makers, where we suppose the same order restrictions of the implausibility measures. Given that we are working in the Anscombe and Aumann framework, and despite the fact that the preferences here are not biseparable, this is similar in spirit to the definition in Ghirardato and Marinacci (2002).}

**Definition 8.** Let $\succeq^1$ and $\succeq^2$ be represented by the same utility and implausibility measures which have the same order restriction on probabilities. Then $\succeq^1$ is more ambiguity averse than $\succeq^2$ if, for any $f \in \mathcal{A}$ and $c \in \Delta(X)$, if $f \succeq^1 c$ then $f \succeq^2 c$.

Although the specification that the utilities are the same is superfluous in this definition, because it is implied by the condition (applied to constant acts), it is retained to facilitate comparison with Definition 7. Below, we shall show that these two notions of aversion to ambiguity or lack of confidence are in fact equivalent.

Before, note that in this framework there is a natural notion of the risky option corresponding to a given ambiguous act. Suppose that the preferences satisfy the axioms of Theorem 1 and that the centre probability is $p$. For any act $f \in \mathcal{A}$, we define the risk-equivalent, $c_f$, as follows: $c_f(s)(x) = p(f^{-1}(x))$, for all $s \in S$, $x \in X$. $c_f$ is the constant act yielding the risky or “objective” lottery which is the same as that generated by $f$, if the centre probability function (ie. the agent’s best estimate of the probabilities) is used. We can use this observation to define the following notion of ambiguity premium.

**Definition 9.** Suppose that the axioms of Theorem 1 hold and let $f$ be an act in $\mathcal{A}$. The ambiguity premium of $f$ is the unique solution of the following equivalence:

$$f \sim c_{f+(-\nu)}$$

We will henceforth write $\nu(f)$ for the ambiguity premium of $f$.

This definition is analogous to the classical definition of risk premium (Pratt, 1964). The risk premium is the amount one is willing to remove from the actuarial value of a risky option to obtain a sure amount indifferent to the risky option; analogously, the ambiguity premium is the amount one is willing to remove from the risky lottery corresponding to the ambiguous act to obtain a risky lottery indifferent to the ambiguous act. A significant difference is that the risk premium is measured in monetary units, whereas the ambiguity premium, as defined, is measured in units of utility. This difference is deliberate: given that ambiguity aversion comes “on top of” risk aversion – witnessed by the fact that comparisons of ambiguity attitude suppose or imply that risk attitudes coincide (Klibanoff et al., 2005; Epstein, 1999; Ghirardato and Marinacci, 2002) – it seems natural to define the premium on a scale where the risk aversion has
already been factored out. However, the difference is not necessary: with appropriate modifications of the definition above, a definition of ambiguity premium in monetary units can be given.

It turns out that comparative aversion to lack of confidence is nicely characterised in terms of the relationship between the implausibility measures. Furthermore, it is equivalent to the comparative ambiguity aversion (in the sense of Definition 8) and to the natural ordering of ambiguity premiums.

**Theorem 2.** Let $\preceq^1$ and $\preceq^2$ be represented by the same utility $u$ and implausibility measures $d_1$ and $d_2$ which have the same order restriction $\leq$. The following are equivalent:

(i) $\preceq^1$ is more averse to lack of confidence than $\preceq^2$

(ii) $\preceq^1$ is more ambiguity averse than $\preceq^2$

(iii) for all $f \in A$, $\nu_1(f) \geq \nu_2(f)$

(iv) $d_1(p) \leq d_2(p)$ for all $p \in \Delta(\Sigma)$.

Note that in the theorem the utilities have to be the same, not just the same up to multiplicative transform: this is required because the calibration of $d_1$ and $d_2$ depends on the utilities. By contrast, in Definitions 7 and 8, the zeroed utilities only need to be the same up to multiplicative transform. Of course, if the utilities are the same up to multiplicative transform, they can always be put into a form such that they are the same.

4 Discussion

4.1 Properties and Assumptions

4.1.1 Outcome space The set of outcomes was assumed to be the reals, for the simplest motivation of the ideas is in terms of monetary payoffs. However, this assumption is stronger than strictly needed: any outcome space supporting an unbounded utility function could be used. (The certainty equivalents, which are needed for the result, are provided by the mixture structure on the consequence space.) Moreover, unboundedness is only really required for Definition 3, in its current form. It may be possible to do without unboundedness, given sufficient modifications to the definition and the theorem.

4.1.2 Monotonicity on outcomes and unboundedness of outcomes As above, the monotonicity on outcomes condition in axiom A1 was posed largely for reasons of ease, and to focus the presentation on important matters. Naturally, a more subtle version of Definition 3, with appropriate modifications of its uses, would be required were this assumption to be dropped. By contrast, the unboundedness in outcomes condition is required to ensure certainty equivalents. Note that it is a weakening of the traditional monotonicity or dominance axiom (Appendix A), which could be used in its place. Given the interpretation of the outcomes as monetary, both monotonicity on outcomes and unboundedness of outcomes are standard assumptions in economics.
4.1.3 Reference point  The proposed representation assumes a special “zero” point on the utility scale, in particular via the assumption of maximality of the $\Lambda$ in the Centering axiom (A6). This assumption is not particularly controversial empirically. Since the work of Kahneman and Tversky (1979), it has been widely accepted that reference points may play an important role in decisions: one seems to take account of the possible gains and losses with respect to a reference point, rather than the possible total wealth values. It need not be assumed that the reference point is the zero monetary gain or loss. Indeed, given that the stakes are defined in terms of the worst-case loss, the reference point is perhaps better thought of as a “satisfaction” level, perhaps monetarily higher than the agent’s actual wealth, such that below this satisfaction level, the agent is particularly careful with his choices. Technically, theories involving reference points, such as Tversky and Kahneman (1992); Wakker and Tversky (1993), do not elicit the position of the reference point, but assume it to be known; this practice is followed here, with the maximality assumption in A6.

It should nevertheless be emphasised that this assumption is made partly for æsthetic reasons. Were it to be dropped, a representation theorem would be obtained which differs from Theorem 1 in that: firstly, the centre of the implausibility measure (the probability with the lowest implausibility) may take a non-zero implausibility value; and secondly, the utility will be unique up to positive affine transformation, and there will be a corresponding additive factor in the uniqueness clause concerning the implausibility measure. Axiomatically, the only difference would be the dropping of the maximality condition in Axiom A6. The idea that the centre probability has zero implausibility, reinforced as it is by the diagram in Figure 1, is the main motivation for including this calibration of utilities.

4.1.4 Linearity  The proposal made here assumes and exploits a property of preferences which cannot be accounted for by most traditional models; namely the non-linearity of the evaluation functional on utilities. Consider the following pairs of choices: between $(x, A; y, A^c)$ and $(0, S)$, and between $(x \ominus 1000, A; y \ominus 1000, A^c)$ and $(0 \ominus 1000, S)$, where $A$ is an event, $x$, $y$ and 0 are monetary values and $x \ominus 1000$, $y \ominus 1000$ and $0 \ominus 1000$ are monetary values having utilities a thousand units less than the utilities of $x$, $y$ and 0 respectively. Under standard expected utility, and according to the main theories of ambiguity mentionned in the Introduction, if the agent prefers the first bet to the sure amount, then he prefers the bet in the second choice as well. However, intuitively this may not be so: although he may be prepared to take bets where he has “little to lose”, he might be reticent to take bets on the same events whose outcomes are just additive transforms of the initial ones, but where he has “more to lose”. This sort of behaviour seems related to confidence: it is because of his lack of confidence in his probability judgement regarding $A$ that he would not accept to bet at high (utility) stakes. Although little experimental work has been done to date on the differences in attitudes to ambiguity at different payoffs, note that in the domain of risk, it is well-known that risk aversion increases as the payoff increases (Holt and Laury, 1986). Respectively: a bet yielding a gain of $x$ dollars if event $A$ occurs and a loss of $x$ if not; and no gain or loss, for sure.

An important exception is the theory proposed by Klibanoff et al. (2005), where this behaviour only occurs under the special case of “constant ambiguity attitude”. See Section 4.4.
We conjecture that a similar phenomenon holds for ambiguity.

4.1.5 Continuity

Given that different acts are evaluated with respect to different sets of probabilities, according to the stakes involved, one might be led to expect that there is a lack of continuity. It is, however, unclear what sort of continuity is lacking. First of all, the traditional continuity axiom on preferences is satisfied (A1). Secondly, the implausibility measure which determines the set of probabilities to be used is continuous. This ensures the following sort of “continuity in stakes”: for any act \( f \), state \( s \) and outcome \( x \), let \( x_s f \) be the act yielding \( x \) on state \( s \) and \( f \) elsewhere, and consider a sequence \( x^i_s f \) such that \( x^i \rightarrow x \). Writing \( U(f) \) for the evaluation of \( f \) as given in (1), \( U(x^i_s f) \rightarrow U(x_s f) \) as \( x^i \rightarrow x \). Although lowering the outcome for a given state may lower the stakes, and thus alter the set of probabilities used in the evaluation, there are no “jumps” as one gradually raises the stakes again.

Another sort of continuity one might want would be “continuity in events”: that a sequence of acts \( x_A f \), with \( x \) the minimal outcome of each act and the sequence \( A \) tending to the empty set, tends to \( f \). However, since the state space of finite, this sequence must be finite, so convergence is not well-defined. The extension of Theorem 1 to infinite state spaces, and investigation of consequences for this sort of continuity, is a topic for future research.

Let us consider one final potential worry of “jumps” in evaluations of acts. Consider \( x_s f \) and \( f \), for some state \( s \) and some outcome \( x \) which is below \( f(s') \) in the preference order, for every state \( s' \). The worry is that, because larger sets of probabilities are used in the evaluation of an act the lower the minimum value of that act, the decision-maker might strictly prefer \( f \) to \( x_s f \), even when the probability of \( s \) is zero: but it seems that people should (and perhaps are) indifferent between acts which differ only on states of zero probability.

This apparently unwelcome conclusion is not as troublesome as it may first seem, for the notion of state of probability zero needs to be defined more precisely in contexts where there is ambiguity. Certainly, one does not want the decision-maker to have strict preferences over acts which differ only on states that he is fully confident to have probability zero; that is, on states such that the implausibility of them having non-zero probability is infinite. The non-nullness clause in Definition 4 guarantees that this is indeed the case: the decision-maker is indifferent between \( x_s f \) and \( f \) for null \( s \).

One could ask for more: that the decision-maker is indifferent between \( x_s f \) and \( f \) where the centre probability functions assigns zero probability to \( s \) (or more generally, where the implausibility of the probability being non-zero is greater than a certain non-infinite degree). This is not implied by the representation in Theorem 1; for this, Definition 4 would have to be modified. Such a modification would yield a “riskier” decision rule than the one proposed here. Suppose, for example, that you think that there is zero probability that the Fed will go bankrupt in the next five years, but you are not fully confident – it is possible but implausible that there is a non-zero probability of bankruptcy. Consider the choice between a particular investment plan which is insured against the collapse of the Fed and the same investment plan where all the money is lost if the Fed collapses. According to the modified decision rule just mentioned, you

---

10Thanks to Itzhak Gilboa for raising this point.
will use the same sets of probability functions in the evaluation of the two options: for certain pairs of investment plans, all of these functions will give zero probability to the collapse of the Fed, so you will be indifferent between the two options. By contrast, according to the decision rule axiomatised in Theorem 1, you will use a larger set of probability functions in the evaluation of the second option (the one without insurance). This set might contain probability functions giving non-zero probability to the collapse of the Fed, even though the set used in the evaluation of the first option does not: hence you will, rightly it seems, strictly prefer the former option to the latter. Evidently, the rule in Theorem 1 is "safer" than the modification just mooted; there are at least *prima facie* reasons for not considering the fact that it may treat functions which differ on states which have zero probability according to the centre probability function differently as a disadvantage.

4.2 Variants

4.2.1 *Expected utility and maxmin* The procedure adopted in Section 2.2 employs well-known decision theories to represent the preferences over acts with the same stakes: namely, expected utility and maxmin expected utility (MMEU). These are natural choices. If one thinks of confidence in terms of nested "spheres" of probability functions (as in Figure 1), with each sphere containing all the functions closer to the centre than a particular distance, and if one accepts that the decision-maker is more confident that the "right" probability function is in a sphere the larger the sphere, then it is natural to work with sets of probabilities as in MMEU.

However, variants can easily be imagined. Most naturally, one could replace the minmax rule with any rule which operates with (possibly different) sets of probability functions, such as $\alpha$-maxmin (Ghirardato et al., 2004), for example. This would involve a modification of the MMEU axiom (A3).

Similarly, one could replace the expected utility "centre": instead of there being a single probability function, there may be a set of probability functions, and one could choose according to whatever rule one has decided on (maxmin in this case). Such a representation is immediately obtained if one drops the Centering axiom (A6); see also Section 4.1.3.

4.2.2 *Stakes* $\sigma$ The motivation for the notion of "stakes" used (Definition 4) is obvious, if not beyond controversy: the agent judges what is at stake by the worst possible outcome of the choice – in other words, by the most he has to lose. Challenges to this notion of stakes do not affect the theorem proposed, for relatively little rests on the precise definition of the "stakes" function $\sigma$. Similar representation theorems can be proved for different functions $\sigma$; indeed, the theorems are formulated in an almost identical manner to that given.

There are naturally a myriad of stakes functions which could be proposed or discussed: for example, the difference between the maximum and minimum values of the acts, the probability that the act takes values below a certain reference point (calculated with the centre probability) or the expected utility of values below the reference point (calculated with the centre probability). A full discussion of all the options is beyond the scope of this paper. A sense of the sorts of issues is given in Section 4.1.5, where
the use of the centre probability leads to a less “safe” decision rule. Rather, we shall just stress two points.

Firstly, although the notion of stakes used here goes best with additive shifts of acts (Definition 3), in particular via the definition of preferences “as if the stakes were $r$” (Definition 5), other notions of stakes might go best with multiplicative shifts (where addition in Definition 3 is replaced with multiplication). Secondly, different notions of stakes imply different properties of the preference order. For example, nowhere is monotonicity (also called dominance) posed as a condition on $\preceq$ (see Section 4.1.2). However, with the definition of stakes proposed above, it is easy to see that $\preceq$ satisfies dominance. However, with other stakes functions, although an analogue of Theorem 1 may hold, the resulting preference may not satisfy dominance; this is the case for example if the stakes function takes the difference between the maximum and minimum values of the act.

4.3 Ambiguity attitude

4.3.1 Other ambiguity concepts In Section 3, only a comparative notion of aversion to lack of confidence was defined, which was intuitively and technically seen to correspond to a sort of ambiguity aversion. In the model proposed above, aversion to ambiguity is built-in (and so cannot be defined), in much the same way as it is built into the maxmin expected utility model, and for the same reason. This observation immediately suggests the following way of extending the definition to cover cases of neutrality to ambiguity (or lack of confidence) and of ambiguity seeking, and to distinguish between these cases and ambiguity aversion. Ambiguity aversion occurs when the agent chooses according to the rule in (1). Ambiguity neutrality occurs when he always chooses according to the expected utility with the centre probability (the confidence, represented by the implausibility measure, plays no role). He is ambiguity seeking if he chooses according to a rule which is as (1) except that minima are replaced by maxima (cf. the maxmax rule swaps maxmin ambiguity aversion for ambiguity seeking). Just as for ambiguity aversion, one can define the relation “more ambiguity seeking” as in Definition 7, but for agents with ambiguity seeking preferences (ie. using maxima rather than minima).

4.3.2 Separating doxastic attitude and ambiguity attitude Decision theorists often want their models to separate out the utility elements in decision making (which, under the classical theory, capture risk attitude) from the elements representing the agent’s beliefs, or more generally his doxastic state, and, in the case of models of ambiguity, to separate each of these from the element prescribing ambiguity attitude (for example Ghirardato et al. (2004) and especially Klibanoff et al. (2005)). It is worth seeing how this can be done in the proposed model.

Under the proposed model, the doxastic state of the agent (very roughly, his “beliefs”) is comprised not only of his best estimate of the probabilities, but of his confidence in these estimates. However, as suggested at the beginning of Section 3, the

11 Suppose, $g(s) \preceq f(s)$ for all $s$. Then $\sigma(g) \succeq \sigma(f)$; since there is dominance, $\min_{\sigma(p) \leq \sigma(g)} \sum u(g(s))p(s) \leq \min_{\sigma(p) \leq \sigma(g)} \sum u(f(s))p(s) \leq \min_{\sigma(p) \leq \sigma(f)} \sum u(f(s))p(s)$, since $\sigma(g) \succeq \sigma(f)$. 
confidence is adequately represented by an order on the set of probability functions: it is the order restriction of the implausibility measure (Definition 2), rather than the implausibility measure per se, which represents the agent's doxastic states. By contrast, the ambiguity attitude of the agent is properly captured by what the implausibility measure adds to its order restriction: it relates the agent's confidence, represented by the order on the space probability functions, to choices, and in particular to the stakes required for different probability functions to play a role in the evaluation of a prospect.

A simple reformulation of representation (1) will make it clear that this is indeed a separation of beliefs and ambiguity attitude. Given an order restriction \( \leq \) on the space of probability functions \( \Delta(\Sigma) \), define \( \Xi \) to be the set of subsets of \( \Delta(\Sigma) \) such that, for any \( p \in \Delta(\Sigma) \), there is a unique \( C \in \Xi \) with \( q \in C \) iff \( q \leq p \). It is easy to see that one can generate the order \( \leq \) given \( \Xi \): \( \Xi \) is thus an equivalent representation of the agent's doxastic state. Now, it is clear that an implausibility measure \( D \) whose order restriction is \( \leq \) generates a function \( D : \Re_{\leq 0} \cup \{\infty\} \to \Xi \) such that \( D(r) = C \) iff, for all \( p \in \Delta(\Sigma) \), \( d(p) \leq r \) iff \( p \in C \) (in fact, \( D(r) = d^{-1}([0, r]) \)). Conversely, from such a function, one can retrieve the implausibility measure: \( D \) represents precisely what the implausibility measure adds above its order restriction. Of course, representation (1) can be reformulated in terms of \( \Xi \) and \( D \) as follows: for all \( f, g \in \mathcal{A}, f \preceq g \) iff

\[
\min_{p \in D(\sigma(f)))} \sum_{s \in S} u(f(s)).p(s) \leq \min_{p \in D(\sigma(g)))} \sum_{s \in S} u(g(s)).p(s)
\]

Under this reformulation, the separation of ambiguity attitude and doxastic state is clear: the doxastic state is the range of \( D \), the ambiguity attitude is the function \( D \) itself.

4.3.3 A co-efficient of ambiguity aversion In fact, not only is the \( D \) introduced above the appropriate representation of the ambiguity attitude, it can be thought of as an analogue of the Arrow-Pratt co-efficient of risk-aversion. First of all, it is an immediate consequence of Theorem 2 that \( \preceq^1 \) is more ambiguity averse than \( \preceq^2 \) if and only if \( D_1(r) \succeq D_2(r) \) for all \( r \in \Re_{\leq 0} \). Secondly, the ambiguity premium can be expressed as a function of \( D \): it is not too difficult to see that, for any \( f \in \mathcal{A}, \nu(f) = \sum_{s \in S} u(f(s)).p^c(s) - \min_{p \in D(\sigma(f)))} \sum_{s \in S} u(f(s)).p(s) \), where \( p^c \) is the centre probability. That is, the ambiguity premium can be expressed as a function of \( D \), the utility and the doxastic state \( \Xi \), just as the risk premium can be approximated by a function of the co-efficient of risk aversion.

Of course, the function \( D \) and its relation to the ambiguity premium takes a very different form from the traditional co-efficient of risk aversion and Arrow-Pratt approximation, and indeed from the only other co-efficient of ambiguity aversion proposed in the literature (to our knowledge), namely, that of Klibanoff et al. (2005), which mimics the Arrow-Pratt notion quite closely: their co-efficient is basically the equivalent of the Arrow-Pratt co-efficient, but takes as its argument expected utilities over first-order probabilities (see also Section 4.4). This difference does not necessarily count against the proposal that \( D \) be thought of as a co-efficient of ambiguity aversion.

\[12\]In a companion paper (Hill, 2009), we consider representation results yielding only the order, and compare them to cardinal representation results of the sort obtained here.
First of all, ambiguity differs risk insofar as it involves the question of the decision-makers’ beliefs (and his confidence in them). The proposed co-efficient explicitly brings in this aspect, insofar as it takes values in sets of probabilities (representing the probability judgements of which the agent is confident to a certain degree). Secondly, there are some well-known problems with the Arrow-Pratt analysis of risk aversion – in particular, plausible risk aversion over modest stakes implies implausible risk aversion over large states Rabin (2000) – and it would certainly not be surprising if co-efficients of ambiguity aversion which are modelled on the Arrow-Pratt co-efficient suffer from similar problems. Finally, the proposed co-efficient and analysis of ambiguity has all the advantages of not relying on local methods. Whereas the Arrow-Pratt approximation is but an approximation, which is exact only in special cases and which can be used only under certain conditions, the expression of ambiguity aversion in terms of the proposed co-efficient of ambiguity aversion is always exact.

### 4.4 Related literature

In Gajdos et al. (2008), the Gilboa and Schmeidler (1989) model is extended to incorporate different sets of probability functions in the evaluation of acts. However, there, the appropriate sets of probabilities are derived from an “objective” set of probability functions (modelling information), which is given exogeneously; by contrast, here the variation in the sets of probability functions used is generated endogeneously, by the implausibility measure and the stakes of the act. Gajdos et al. (2008) propose notions of comparative imprecision aversion and imprecision premium, which are difficult to relate to the notions proposed here, because they rely on the exogeneously given sets of probabilities.

The proposal here models the agent’s confidence by a structure on top of the set of probabilities. The most popular model of this type in the literature employs a probability over the set of probabilities and evaluates acts using by the expected value (under the second-order probabilities) of a transform of the expected utilities of the act (under the first-order probabilities) (Klibanoff et al., 2005; Nau, 2006; Seo, 2007). Insofar as both those models and the one proposed here involve a structure on top of the set of probabilities, there are many analogies between the results obtained; however, there are also some significant differences. Technically, of course, a probability measure over a space does not have the same properties as an order, or an implausibility measure: additivity, for one, plays no role here. Conceptually, the interpretations do not seem to correspond either: whereas a probability over probability functions can and often is thought of as a second-order belief, one cannot think of an order over the set of probabilities as a second-order belief, without breaking with the Bayesian principle that beliefs are represented by probabilities. Thus the interpretation as a measure of the degree of confidence. It follows that the proposal made here, contrarily to the one cited above, does not rely on or link into the literature on the reduction axiom and two-stage lotteries (Segal, 1990; Halevy, 2007). For further comparison between the models on detailed points, see Sections 3, 4.1.4, 4.3.2 and 4.3.3.

The intuition involved in the current proposal is not unrelated to that involved in the literature on robustness (for example Hansen and Sargent (2001); for an axiomatisation of a generalisation of this model, see Maccheroni et al. (2006)). There, as here, there
is best estimate probability or model (the “reference probability”, corresponding to our centre probability). There, as here, the idea is that the decision-maker takes into account probabilities (or model specifications) which differ from this probability. There, just as here, there is a single parameter which determines the extent to which other probabilities are involved (the robustness parameter $\theta$ in Hansen and Sargent (2001); $D$ here, see Section 4.3.3). However, there are differences between the proposals. First and foremost, the functional forms are different. For one, Hansen and Sargent take the infimum over the whole space of probabilities, whereas we use restricted but variable sets of probabilities. Accordingly, they do not relate the set of probabilities involved in a decision to what is at stake, but determine their “involvement” entirely from the robustness parameter and the relative entropy. Finally, whereas their model is dynamic, we have only presented a static model here (though see below).

To our knowledge, the only suggestion involving a similar structure on the set of probabilities is by Gärdenfors and Sahlin (1982). They introduce a measure of “epistemic reliability” on the set of probabilities, and suggest that, in any decision, the decision-maker should chose a level of epistemic reliability, and then make his choice according to the maximin decision rule over the set of probabilities whose reliability is greater than this level. The notion of “epistemic reliability” appears to be the dual of the notion of “implausibility measure” used here, and the use of the minmax rule over a variable set of probability functions is an important common point between this paper and Gärdenfors and Sahlin’s. However, they do not specify how the desired level of epistemic reliability is to be chosen, and they do not relate it to the notion of stakes as we do; consequently, they do not propose a representation theorem nor any notion of comparative ambiguity aversion. Furthermore, in a choice over a set of acts, Gärdenfors and Sahlin use the same set of probabilities to evaluate all acts, whereas under (1), different sets of probabilities will be used if the acts have different stakes.

Finally, let us note that the order restriction of the implausibility measure resembles the order on possible worlds used in the theory of belief revision (Gärdenfors, 1988), and which is supposed to represent the “epistemic entrenchment” of particular beliefs – approximately, how unwilling one is to surrender the belief. Of course, the notion of confidence and that of epistemic entrenchment are related: the less willing one is to surrender a belief, the more confidence one has in it. However, whereas the literature on belief revision considers only the epistemic entrenchment of propositional beliefs (eg. it will rain tomorrow), here we are interested in the confidence in one’s subjective probabilities (eg. the probability that it will rain tomorrow is greater than 0.5). Nevertheless, the relationship is worth noting, because the question of change in epistemic entrenchment has been quite extensively studied by theorists of belief revision (by, among others: Rott (2003); Hill (2008)), and this literature contains suggestions as to how to define dynamic operators for the model of confidence proposed here.

5 Application

We close the paper with a simple portfolio-style example. This is aimed at helping the reader to get concrete idea of the notions introduced in this paper, and at illustrating how the framework may be used in applications. Many writers on ambiguity, not least
Figure 2: The options

<table>
<thead>
<tr>
<th></th>
<th>( \omega_1 )</th>
<th>( \omega_2 )</th>
<th>( \omega_3 )</th>
<th>( \omega_4 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( H )</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>( A )</td>
<td>4</td>
<td>1</td>
<td>4</td>
<td>1</td>
</tr>
</tbody>
</table>

Gajdos et al. (2008) and Klibanoff et al. (2005), provide examples of applications of their models, noting the possible importance of ambiguity for understanding phenomena such as the equity premium puzzle or home bias (see Klibanoff et al. (2005) for further references). We shall consider a simple version of the latter phenomenon, to which our model seems particularly appropriate, given the intuition that investors are less confident in their judgements regarding foreign markets than in those concerning the home market.

In the example, the agent has one unit of wealth which he must invest and there are two types of stocks on offer (H: in the home market; A: in the foreign market), the returns of which are given in Figure 2. The problem is to decide what proportion \( a \) of his budget to invest in the home stock (he will invest the rest in the foreign stock). The agent has beliefs and confidence represented by the following order on the set of probability functions. He is fully confident that the event that the home market stock yields returns (\( \{\omega_1, \omega_2\} \)) is independent from the event that the foreign market stock yields returns (\( \{\omega_1, \omega_3\} \)), so each probability function which he considers not totally implausible can be associated with a unique pair \( (p_1, p_2) \), where \( p_1 \) is a probability of \( \{\omega_1, \omega_2\} \) and \( p_2 \) is a probability of \( \{\omega_1, \omega_3\} \). Under his best estimate, there is a 0.7 probability that the home stocks will yield returns and a 0.8 probability that the foreign ones will yield returns: the centre probability is \((0.7, 0.8)\). The other probabilities are farther from these, the less implausible they are: for any \( c, d \in \mathbb{R}_{>0} \), if \( d \geq c \), then \((0.7 + d, 0.8 + 3.d) \sim (0.7 - d, 0.8 - 3.d) \geq (0.7 - c, 0.8 - 3.c) \sim (0.7 + c, 0.8 + 3.c)\).

The difference by a factor of 3 represents the fact that the agent is less confident of his judgement about the foreign markets: the implausibility that the probability of returns on the home stock is 0.05 lower than his best estimate is the same as the implausibility that the probability of returns in the foreign stock is 0.15 lower than his best estimate.

Note that each probability function is uniquely characterised by the difference between the probability of yield of the home stocks and his best estimate (the \( c \) and \( d \) above).

We take a utility function displaying constant absolute risk aversion, which is normalised so that \( u(5) = 0 \) and \( u(0) = -5 \); ie.

\[
u(x) = \frac{5}{1 - e^{-5.\gamma}} (e^{-5.\gamma} - e^{-\gamma.x})
\]

where \( \gamma \) is the coefficient of absolute risk aversion.

Note that in this simple example, the stakes of all acts are the same, so the contribution of ambiguity aversion is resumed in the set of probabilities corresponding to these stakes. As noted above, this set is characterised by a positive real number \( \delta \) (it is the set of probabilities \( \{(0.7 + c, 0.8 + 3.c), (0.7 - c, 0.8 - 3.c)\mid c \leq \delta\} \) for some \( \delta \).
To consider the effects of different degrees of ambiguity aversion on choice, it suffices to consider the choices for different values of $\delta$.

**Figure 3: Optimal fraction $a$ of budget allocated to H**

<table>
<thead>
<tr>
<th>$\gamma$</th>
<th>$a$</th>
<th>$\delta$</th>
<th>$a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>.236584</td>
<td>0</td>
<td>.345260</td>
</tr>
<tr>
<td>1</td>
<td>.454473</td>
<td>0.02</td>
<td>.405756</td>
</tr>
<tr>
<td>2</td>
<td>.543300</td>
<td>0.04</td>
<td>.454473</td>
</tr>
<tr>
<td>5</td>
<td>.580084</td>
<td>0.06</td>
<td>.496366</td>
</tr>
<tr>
<td>10</td>
<td>.590081</td>
<td>0.08</td>
<td>.534332</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.10</td>
<td>.570350</td>
</tr>
</tbody>
</table>

Once these assumptions have been made, it is straightforward to perform comparative statics on the co-efficient of risk aversion ($\gamma$) and the parameter representing ambiguity aversion ($\delta$). Typical examples of the effects of risk aversion on investment in H for a given value of the ambiguity aversion parameter, and of the effects of ambiguity aversion for a given value of the co-efficient of risk aversion are given in Figure 3. As risk aversion goes up, the agent invests more in the home stock. However, the same effect occurs if risk aversion is held fixed, and ambiguity aversion is increased. Recall that as the parameter $\delta$ increases, the agent is using probability functions which are more implausible with respect to both stocks. However, the set of possible probability values for the event that the foreign stocks yield returns is expanding faster than the corresponding set for the event that the home stock yields returns: thus the movement away from the foreign stock. Of course, this is the behaviour that corresponds to home bias.\(^{13}\)

In fact, the strength of the effect of the agent’s ambiguity aversion depends on his risk aversion. At low or moderate levels of risk aversion, differences in ambiguity aversion have important effects on the investment choice. By contrast, when the risk aversion is high, it dominates the choice, and differences in ambiguity aversion have only modest effects. This can be seen in Figure 4, which plots the optimum investment in H against ambiguity aversion, for different levels of risk aversion.\(^{13}\)

\(^{13}\)Note that the foreign stock is more attractive than the home one in all aspects except ambiguity; so the optimum portfolio for an ambiguity neutral (and risk averse) agent would be to diversify, but with a larger investment in the foreign stock.
Figure 4: Graph of optional fraction invested in H (a, on the y-axis) against ambiguity aversion (d, on the x-axis). The lines correspond to risk aversion 0.5, 1, 2, 5, 10 respectively (0.5: the line cutting the y-axis at the lowest point, to 10: the line cutting the y-axis at the highest point).

6 Conclusion

This paper proposes a model of the decision-maker’s confidence in his probability judgements in terms of an implausibility measure – a real-valued function on the set of probability functions. A decision rule according to which the decision-maker evaluates acts using sets of probability functions which depend on what is at stake in the choice of the act is axiomatised. The framework proposed yields a natural notion of comparative aversion to lack of confidence and of ambiguity premium. It is shown that these notions are equivalent and can be characterised in terms of the implausibility measure representing the agent’s confidence; moreover they are equivalent to a notion of ambiguity aversion which is close to those defined elsewhere (Klibanoff et al., 2005; Ghirardato and Marinacci, 2002). Finally, it is shown, via an example, how to apply this model to portfolio problems.
A Properties of preferences

Here we recall the definitions of some basic properties on preference orders which were used in Section 2.2. For further discussion, see for example Anscombe and Aumann (1963); Gilboa and Schmeidler (1989); Gilboa et al. (2008). Let $\preceq$ be a preference order on $\mathcal{A}$; then:

**Weak order** $\preceq$ satisfies weak order if (1) for all $f$, $g$ and $h$ in $\mathcal{A}$, if $f \preceq g$ and $g \preceq h$, then $f \preceq h$ and (2) for all $f$, $g$ in $\mathcal{A}$, $f \preceq g$ or $g \preceq f$.

**Non triviality** $\preceq$ satisfies non triviality if there are elements $f$ and $g$ in $\mathcal{A}$ such that $f \preceq g$.

**Independence** $\preceq$ satisfies independence if, for all $f$, $g$ and $h$ in $\mathcal{A}$, and for all $a \in \mathbb{R}$, $0 < a < 1$, if $f \preceq g$ then $af + (1-a)h \preceq ag + (1-a)h$.

**C-independence** $\preceq$ satisfies C-independence if, for all $f$, $g$ in $\mathcal{A}$, all constant acts $h \in \Delta(X)$, and all $a \in \mathbb{R}$, $0 < a < 1$, if $f \preceq g$ then $af + (1-a)h \preceq ag + (1-a)h$.

**Continuity** $\preceq$ satisfies continuity if, for all $f$, $g$ and $h$ in $\mathcal{A}$, the sets $\{a \in [0,1] | af + (1-a)h \preceq g\}$ and $\{a \in [0,1] | af + (1-a)h \succeq g\}$ are closed in $[0,1]$.

**Monotonicity** $\preceq$ satisfies monotonicity if, for any $f,g$ in $\mathcal{A}$, if $f(s) \preceq g(s)$ for all $s$ in $\mathcal{S}$, then $f \preceq g$.

**Uncertainty Aversion** $\preceq$ satisfies uncertainty aversion if, for all $f$, $g$ in $\mathcal{A}$, and all $a \in \mathbb{R}$, $0 < a < 1$, if $f \sim g$ then $af + (1-a)g \succeq f$.

B Proofs

*Proof of Theorem 1.* By A1 and A2, there is a utility $u$, unique up to positive affine transformation, representing the restriction of $\preceq$ to the constant acts. Given the uniqueness, we can suppose that $u$ is zeroed; ie. that it is picked so that $u(\Lambda) = 0$.

For any $r$, from A1 it follows that $\preceq_r$ satisfies non triviality, weak order and continuity, which, in conjunction with A3, implies that there is a closed and convex set of probability functions $\mathcal{C}_r$ and a utility $u_r$ representing $\preceq_r$ by the MMEU formula (Gilboa and Schmeidler, 1989; Ghirardato et al., 2004). By A4, the $u_r$ are a positive affine transformations of $u$ for all $r$, and $\mathcal{C}_r \subseteq \mathcal{C}_r'$ for all $r \leq r'$ such that there exists $s \geq r$ or $r'$ with $\preceq_s \neq \preceq_r$ or $\preceq_r' \neq \preceq_r$ (Ghirardato et al., 2004, Proposition 6). Define $d(p)$ as follows: for any $p \in \Delta(\mathcal{S})$, $d(p) = \inf_{p \in \mathcal{C}_r} r$ (where the convention is taken that the infimum is $\infty$ if the set is empty). By Lemma 1 (below) and A4, $\{r | p \in \mathcal{C}_r\}$ is closed, so $p \in \mathcal{C}_{d(p)}$. By the stipulation that the utility function is zeroed and axiom A6, $d^{-1}(0)$ contains just one element.
Now let us show that \( d \) is continuous. It suffices to show that, for any \( r < \sup_{p \in \mathcal{P}} d(p) \neq \infty d(p) \), the inverse image under \( d \) of \( [0, r] \) is closed and the inverse image of \( [0, r] \) is open. The first part is immediate: \( d^{-1}([0, r]) = C_r \), which is closed. As for the second part, it suffices to show that if \( p \in \Delta(X) \) is on the boundary of \( C_r \), then \( d(p) = r \) (by the contrapositive, it follows that, if \( d(p) < r \), then \( d(p) < s < r \) for some \( s \), so \( p \in \text{int}(C_s) \subset C_s \subset C_r \); hence it is in an open ball in \( d^{-1}([0, r]) \)).

Consider a probability function \( p \) on the boundary of \( C_r \). Recall that we are working in the space \( ba(S) \) of finitely additive real-valued set functions on \( S \), under the weak* topology (this notation is borrowed from Dunford and Schwartz (1958, Definition IV.5.2)). By a separation theorem (Aliprantis and Border, 2007, 5.67), there is a continuous linear functional \( \phi \) on \( ba(S) \) and \( \alpha \in \mathbb{R} \) such that \( \phi(q) \leq \alpha \leq \phi(q) \) for all \( q \in C_r \).

Since \( S \) is finite (so \( ba(S) \) and \( B(S) \) are finite-dimensional, \( ba(S)^* \) is isometrically isomorphic to \( B(S) \), the space of bounded real-valued functions on \( S \) (Dunford and Schwartz, 1958, IV.13.5 and IV.5.3); hence there is a real-valued function \( F \) such that \( \phi(q) = \sum F(s)q(s) \) for any \( q \in ba(S) \). Consider the act \( f \) such that \( u(f(s)) = F(s) \) for all \( s \in S \), and the constant act \( \delta_\alpha \in X \) such that \( u(\delta_\alpha) = \alpha \). By construction and A4, \( \delta_\alpha \succeq_r f \) for all \( r' \leq r \), and \( \delta_\alpha \succeq_{d(p)} f \); by A5, it follows that \( d(p) \geq r \). However, \( p \in C_r \), so \( d(p) = r \), as required.

The other properties of the representation are straightforward to check. Note in particular that, because every act has a certainty equivalent (thanks to A1) and constant acts are evaluated the same way by all \( \preceq_r \), for every \( r \), the representations of the \( \preceq_r \) extend immediately to a representation of \( \preceq \).

Uniqueness comes from the uniqueness of the EU and MMEU results.

\[ \square \]

**Lemma 1.** Under the axioms in Theorem 1 (A1–A6), for any \( r \in \mathbb{R} \), \( C_r = \bigcap_{r' > r} C_{r'} \).

**Proof.** If there is no \( s \in \mathbb{R} \) such that \( s \geq r \) and \( \zeta_{r'}^* \neq \zeta_r^* \), then \( \zeta_{r'}^* = \zeta_r^* \) for all \( r' > r \) and the Lemma is trivially true. Suppose that this is not the case. By A4, \( \zeta_{r'}^* \leq \zeta_r^* \) for all \( r' > r \) so \( C_r \subseteq \bigcup_{r' > r} C_{r'} \) for all \( r \). Suppose, for reductio, that \( C_r \subseteq \bigcap_{r' > r} C_{r'} \), so that there exists a point (probability) \( p \notin C_r \), but \( p \in \bigcap_{r' > r} C_{r'} \). Under the weak* topology on \( ba(S) \), \( ba(S) \) is locally convex (Dunford and Schwartz, 1958, §V.3), and the set of finitely additive probability functions is compact, so \( C_r \) is a closed, compact, convex set. By a separation theorem (Dunford and Schwartz, 1958, V.3.9), there is a continuous linear functional \( \phi \) on \( ba(S) \) and \( \alpha \in \mathbb{R} \) such that \( \phi(p) \leq \alpha < \phi(q) \) for all \( q \in C_r \). As in the proof of Theorem 1, there is a real-valued function \( F \) such that \( \phi(q) = \sum F(s)q(s) \) for any \( q \in ba(S) \). Consider the act \( f \) such that \( u(f(s)) = F(s) \) for all \( s \in S \) and the constant act \( \delta_\alpha \in X \) such that \( u(\delta_\alpha) = \alpha \). By construction and A4, \( \delta_\alpha \succeq_r f \) for all \( r' > r \) but \( \delta_\alpha \nless_r f \), contradicting A5; hence \( C_r = \bigcap_{r' > r} C_{r'} \).

\[ \square \]

**Proof of Theorem 2.** Let the assumptions of the theorem be satisfied.

(i) implies (ii). For some \( f \in \mathcal{A} \) and \( c \in \Delta(X) \), suppose that \( f \succeq_r 1 \). By (1), \( f \succeq_r c \) for all \( r \leq \sigma_1(f) \). By (i), it follows that \( f \succeq_r c \) for all \( r \leq \sigma_1(f) \). But since the preferences are represented using the same utilities, \( \sigma_1(f) = \sigma_2(f) \), and \( f \succeq c \).

(ii) implies (iii). For any \( f \in \mathcal{A} \), \( f \succeq c + (-\nu_1(f)) \), so by (ii), \( f \succeq c + (-\nu_1(f)) \). So \( \nu_2(f) \leq \nu_1(f) \).
(iii) implies (ii). Evident.
(ii) implies (iv). If \( d_1(p) = 0 \), there is nothing to check, and if \( d_1(p) = \infty \), then \( p \) appears at the top of the order restriction, so \( d_2(p) = \infty \). Consider the remaining case, where \( d_1(p) \) is a non-zero real number. For any \( f \in A \) and \( c \in \Delta(X) \) such that \( f \preceq_{d_1(p)} c \), it follows from (1) that \( f \preceq_{1} c \) for any \( r \leq d_1(p) \). By Definition 2, this means that \( [f + (\sigma_1(f) - r)]_1 + (r - \sigma_1(f)) \preceq_{1} c \) for every \( r \leq d_1(p) \), which is equivalent to \( f + (\sigma_1(f) - r) \preceq_{1} c + (\sigma_1(f) - r) \) for every \( r \leq d_1(p) \). So, by (ii) and the fact that \( \sigma_1(f) = \sigma_2(f) \), \( f \preceq_{2} c \) for all \( r \leq d_1(p) \). By representation (1), and the fact that the preferences are represented with the same utility, it follows that \( d_2(p) \geq d_1(p) \).

(iv) implies (i). Consider any \( f, g \in A \) and \( \rho \in \mathbb{R} \) with \( f \preceq_{1} g \) for all \( r \leq \rho \). By representation (1), this assumption implies that for any \( r \leq \rho \), MMEU with \( \{p \in \Delta(\Sigma) : d_1(p) \leq r\} \) yields a weak preference for \( g \) over \( f \). However, since the representations share the same order restriction on \( \Delta(\Sigma) \) and \( d_1(p) \leq d_2(p) \) for all \( p \in \Delta(\Sigma) \), for any real number \( s \), there is a number \( s' \leq s \) such that \( \{p \in \Delta(\Sigma) : d_2(p) \leq s'\} = \{p \in \Delta(\Sigma) : d_1(p) \leq s\} \). It follows that each of sets \( \{p \in \Delta(\Sigma) : d_2(p) \leq r\} \) for all \( r \leq \rho \) equals some \( \{p \in \Delta(\Sigma) : d_1(p) \leq r'\} \), where \( r' \leq \rho \). By representation (1), it follows that \( f \preceq_{2} g \) for all \( r \leq \rho \).

References


