Theoretical Economics 4 (2009), 199–225 1555-7561/20090199

Choice deferral and ambiguity aversion

Igor Kopylov
Department of Economics and Institute for Mathematical Behavioral Sciences,
University of California, Irvine

When confronted with uncertain prospects, people often exhibit both choice deferral and Ellsberg-type ambiguity aversion. This paper obtains a joint representation for these behavioral phenomena. The decision maker as portrayed by my model is willing to choose an uncertain prospect f over g rather than to defer this choice if and only if the expected utility of f is greater that or equal to the expected utility of g for every probability measure in a convex and closed set Δ. This set is interpreted as a collection of the decision maker's possible future beliefs. When choices cannot be deferred, the decision maker evaluates every uncertain prospect via an ε-mixture of the least favorable element in the set Δ and her current probabilistic belief $p \in \Delta$. All components of my model are derived from observable preferences in an essentially unique way.

Keywords. Choice deferral, ambiguity aversion, epsilon contamination, multiple priors model, subjective probability, Ellsberg paradox.

JEL classification. D81, D83.

1. Introduction

People often defer choices among uncertain prospects until they get thoroughly informed about the process through which uncertainty will be resolved. For example, business managers postpone selling a new product until they run safety tests, investors wait to allocate their portfolios until they consult with independent experts, advantage gamblers postpone wagering until they learn the mathematical odds of winning. Intuitively, such decisions are deferred because people are uncertain about their future beliefs and hence about their future preferences over feasible alternatives. This intuition is consistent with empirical studies of choice deferral in psychology and marketing (e.g. Tversky and Shafir 1992 and Dhar 1997). In decision theory, uncertainty about tastes has been used to explain preference for flexibility as in Koopmans (1964), Kreps (1979), and Dekel et al. (2001).

When choices cannot be postponed and objective information is scarce, people often exhibit *ambiguity aversion*. To illustrate, recall Ellsberg's (1961) famous paradox where the decision maker is told only that (i) a ball will be drawn randomly from an

Igor Kopylov: ikopylov@uci.edu
Discussions with Larry Epstein, Peter Wakker, and Bart Lipman were essential for this paper. I thank Kyoungwon Seo, Jean-Claude Falmagne, Massimo Marinacci, Stephen Morris, and all anonymous referees for their comments.

urn that contains 90 balls of three possible colors (red, green, and blue), and (ii) the number of red balls in the urn is 30. Then it is typical to bet on the event $\{R\}$ rather than on $\{B\}$ because the decision maker knows the objective probability of $\{R\}$ to be $\frac{1}{3}$, but does not know the objective probability of $\{B\}$. Analogously, it is typical to bet on $\{B,G\}$ rather than on $\{R,G\}$. These betting preferences cannot be represented by any probability measure p because the inequalities $p(R)>p(G)$ and $p(G)+p(B)>p(R)+p(B)$ are inconsistent.

Ellsberg's setting can be adapted to illustrate choice deferral as well. To do so, suppose that decisions can be postponed until the precise composition of the urn is announced. Then the decision maker should defer her choices between bets on $\{R\}$ and on $\{G\}$ and between bets on $\{G,B\}$ and on $\{R,B\}$ because she does not know how she will rank these bets after she learns the composition of the urn. This choice deferral is intuitive even if she is not averse to ambiguity and does not exhibit the preference reversals in the Ellsberg paradox. By contrast, the decision maker should agree to bet on $\{G,B\}$ rather than on $\{R\}$ immediately because she believes that $\{G,B\}$ is more likely than $\{R\}$ and expects to keep the same belief after she learns the composition of the urn.

This paper obtains a joint representation for choice deferral and ambiguity aversion and identifies a behavioral connection between these phenomena. I model two primitive preference relations $\succeq$ and $\succeq_*$ over Anscombe and Aumann's (1963) *acts*—functions that map states of nature into lotteries (i.e. numerical distributions over deterministic prizes.) For any two acts f and g, the preference $f\succeq g$ means that the decision maker is willing to choose f rather than g when no other alternatives are feasible, and the *firm* preference $f\succeq_* g$ means that she is still willing to choose f rather than g when she has the option to postpone this choice. Intuitively, the ranking $f\succeq_* g$ reveals the decision maker's firm belief that she should prefer f to g whenever these acts are available to her. By definition, the ranking $\succeq_*$ is incomplete if choices between some acts f and g are deferred.

My main result (Theorem 1) provides necessary and sufficient conditions for the following pair of representations for the preferences $\succeq$ and $\succeq_*$:

- $\succeq$ is represented by the utility function

$$U(f)=(1-\varepsilon)\int u(f(s))\,dp+\varepsilon\min_{q\in\Delta}\int u(f(s))\,dq, \tag{1}$$

- for all acts $f,g\in\mathscr{H}$,

$$f\succeq_* g \quad\Leftrightarrow\quad \int u(f(s))\,dq\geq\int u(g(s))\,dq \quad \text{for all } q\in\Delta, \tag{2}$$

where u is an expected utility index, $\varepsilon\in[0,1]$, Δ is a convex and closed set of probability measures, and $p\in\Delta$. Moreover, all of these components are derived from $\succeq$ and $\succeq_*$ in an essentially unique way.

Representation (2) is due to Bewley (2002). In my framework, the decision maker as portrayed by (2) expects that she will make her deferred choices via expected utility, and her probabilistic belief will belong to the convex and closed set Δ. Accordingly, she firmly prefers an act f to g if and only if the expected utility of f is greater than or equal to the expected utility of g. To obtain the set Δ from the firm preference $\succeq_*$, I invoke a characterization result due to Ghirardato et al. (2004) (henceforth GMM).

The novel part of my model is representation (1) for the preference $\succeq$. This utility function evaluates every act f via the ε-mixture of the least favorable belief in the set Δ with the probability measure $p \in \Delta$. It is natural to interpret p as the decision maker's ex ante probabilistic belief, and ε as an index of her ambiguity aversion. Theorems 2 and 3 provide some behavioral foundations for this interpretation of p and ε.

Representation (1) can be written in the maxmin expected utility form

$$U(f) = \min_{q \in \Pi} \int u(f(s))\, dq,$$

where the convex and closed set of priors

$$\Pi = (1-\varepsilon)\{p\} + \varepsilon\Delta$$

has the parametric structure of *ε-contamination.*[1] Therefore, representation (1) is a special case of the multiple priors model due to Gilboa and Schmeidler (1989) (henceforth GS).

My key axiom, called Cautious Independence, requires separability:

$$f \succeq g \quad \Rightarrow \quad \alpha f + (1-\alpha)h \succeq \alpha g + (1-\alpha)h$$

for all $\alpha \in [0,1]$ and for all acts f, g, h such that the mixture $\alpha f + (1-\alpha)h$ is a "better hedge" than $\alpha g + (1-\alpha)h$. In particular, this constraint holds when h is constant or when $h = g$, which implies GS's Certainty Independence and Uncertainty Aversion respectively. More broadly, my notion of a "better hedge" is defined in terms of the firm preference $\succeq_*$ and can apply for arbitrary h. Thus, Cautious Independence strengthens the counterpart axioms of the multiple priors model.

Representation (1) has applications in economics, decision theory, and statistics. An important special case is obtained if Δ equals the entire simplex $\mathscr{P}$ of all probability measures on the state space S. In behavioral terms, it means that the decision maker defers her choice between any acts f and g that do not statewise dominate each other. The corresponding utility function (1) takes the form

$$U(f) = (1-\varepsilon)\int u(f(s))\, dp + \varepsilon \min_{s \in S} u(f(s)),$$

and the set of priors Π has the structure

$$\Pi = (1-\varepsilon)\{p\} + \varepsilon\mathscr{P}.$$

[1]The ε-contamination structure was originally proposed by Hodges and Lehmann (1952) in Bayesian analysis (reviewed by Berger 1994). In statistical applications, the parameter ε is interpreted as the amount of error that is deemed possible for the prior p. This interpretation differs from mine because it uses ε to describe the imprecision of a priori knowledge rather than the behavioral effect of this imprecision.

This structure has been used in models of asset pricing (Epstein and Wang 1994), search (Nishimura and Ozaki 2004), and insurance (Carlier et al. 2003.)

Note that representation (1) can be applied in Ellsberg's setting where the natural candidate for Δ is the set of all probability measures q such that $q(R)=\frac{1}{3}$. Ellsberg (1961, pp. 663–669) suggests the functional form (1) as an ad hoc explanation for his paradox. He describes p as an "estimated distribution, which reflects all [subjective] judgements of the relative likelihoods of distributions, including judgements of equal likelihoods," and the parameter $1-\varepsilon$ (ρ in his notation) as a degree of the subjective "confidence in the best estimates of likelihood." My results provide axiomatic foundations for Ellsberg's intuition.

Similarly, one can specify the set Δ exogenously in most experimental studies of ambiguity where only intervals of possible objective probabilities are given to subjects. In an early study of this kind, Becker and Brownson (1964) find some evidence that people put a constant weight ε on the least favorable probabilistic scenario in this set. (This study estimates the average weight ε to be 0.768.) Other experiments (reviewed by Camerer and Weber 1992) produce mixed results.

Yet it should be emphasized that the set Δ in my model can be derived from choice behavior even if the decision maker's knowledge about objective probabilities is not observable, or if the very concept of "objective probabilities" is problematic.

My model is related to several decision-theoretic results. Gilboa (1988) and Jaffray (1988) axiomatize a counterpart of representation (1) for choice of objective lotteries where expected utility is mixed with the worst possible outcome. Eichberger and Kelsey (1999) and Nishimura and Ozaki (2006) characterize ε-contamination with complete ignorance (i.e. $\Delta=\mathscr{P}$) in Anscombe–Aumann's framework. (Nishimura and Ozaki take the parameter ε as a primitive as well.) Gajdos et al. (2008) derive ε-contamination in a different framework that includes variable information sets Δ and incorporates them into objects of choice—it is assumed that the decision maker ranks act-information pairs of the form (f,Δ). These authors interpret the parameter ε as a degree of imprecision aversion, which is common for all information sets. The probability measure p in their model is uniquely determined by Δ and hence does not depend on preference (is not subjective).

The model that is most closely related to mine is due to Gilboa et al. (2008) (henceforth GMMS). These authors characterize a special case of representations (1) and (2) with $\varepsilon=1$. I rely in part on their analysis, but have a different motivation and a more general representation result.

2. Model

2.1 *Preliminaries*

I adopt a version of Anscombe and Aumann's (1963)'s decision framework. A set X of *outcomes*, a set S of *states* of nature, and an algebra $\Sigma\subset 2^S$ of *events* are given. Based on these primitives, define

- the set $\mathscr{L}=\{l,\dots\}$ of all *lotteries*—probability measures on X with finite support

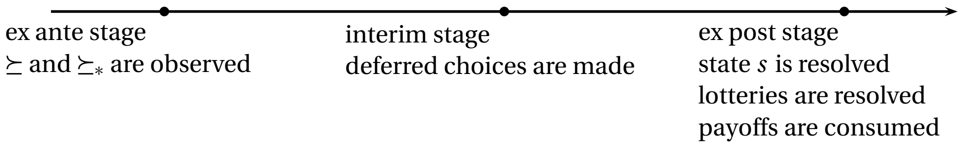

FIGURE 1. The time line.

- the set $\mathscr{U}$ of all expected utility functions on $\mathscr{L}$
- the set $\mathscr{P} = \{p, q, \ldots\}$ of all finitely additive probability measures on (S, Σ) with the weak* topology[2]
- the set $\mathscr{C}$ of all non-empty, convex, closed subsets of $\mathscr{P}$
- the set $\mathscr{H} = \{f, g, \ldots\}$ of all *acts*—Σ-measurable functions $f : S \to \mathscr{L}$ that have a finite range in $\mathscr{L}$.

Endow the set $\mathscr{H}$ with a natural *mixture* operation: for any $f, g \in \mathscr{H}$ and $\alpha \in [0,1]$, let $\alpha f + (1-\alpha) g$ be an act such that for all $s \in S$,

$$[\alpha f + (1-\alpha) g](s) = \alpha f(s) + (1-\alpha) g(s).$$

Identify constant acts with the corresponding lotteries $l \in \mathscr{L}$. Given any lotteries $l, l' \in \mathscr{L}$ and any event $A \in \Sigma$, define a binary act

$$l\,A\,l' = \begin{cases} l & \text{if } s \in A \\ l' & \text{if } s \notin A. \end{cases}$$

Interpret any act $f \in \mathscr{H}$ as a physical action that yields the lottery $f(s)$ after the state s is observed. This lottery is then resolved via an objective randomizing device like a fair coin or a roulette wheel.

For any act $f \in \mathscr{H}$ and any probability measure $q \in \mathscr{P}$, let

$$f(q) = \sum_{l \in \mathscr{L}} l \cdot q\left(\{s : f(s) = l\}\right).$$

This mixture is well-defined because f has finite range. Say that the lottery $f(q)$ is *induced* by f via q.

Suppose that prior to the *ex post* stage when all uncertainty is resolved and payoffs are consumed, there are two time periods, *ex ante* and *interim*, when choices can be made. Let $\succeq$ be the decision maker's ex ante preference over acts in $\mathscr{H}$. As customary, $f \succeq g$ means that she chooses the act f rather than g when no other options are feasible ex ante.

[2] If S is finite, then this topology is Euclidean. In general, a net $\{q_d\}_{d \in D}$ in $\mathscr{P}$ converges to $q \in \mathscr{P}$ in the weak* topology if for any event $A \in \Sigma$, the net $\{q_d(A)\}_{d \in D}$ converges to $q(A)$.

Suppose next that at the interim stage, the decision maker can obtain additional information about the physical process through which uncertainty will be resolved ex post. For example, she may learn some experimental data, experts' opinions, or market odds (e.g. insurance rates, bookmakers' lines). Anticipating the arrival of such information, she may be uncertain ex ante about her interim preferences.

For any acts $f, g \in \mathscr{H}$, let $f \succeq_* g$ if the decision maker is willing to choose f rather than g ex ante even though she has the option to postpone this choice until the interim stage. Intuitively, this behavior reveals her firm belief that she should still prefer f to g at the interim stage regardless of any new information that may arrive by that time. To reflect this intuition, call the relation $\succeq_*$ a *firm* preference. Note that $\succeq_*$ is observed at the ex ante stage, and the decision maker's interim choice behavior is not a primitive of my model.

One concern about the above interpretation of the firm preference $\succeq_*$ is that some people may postpone ex ante choices between acts f and g without even contemplating possible interim preferences between these two acts. They may do so to delay contemplation or to get some intrinsic value of freedom of choice as in Sen (1988). To make the ranking $\succeq_*$ more deliberate, one may adapt the experimental design of Danan and Ziegelmeyer (2008) and set a small monetary cost (such as 10 euro cents) for postponing a choice between acts f and g. This cost should motivate the decision maker to commit to f ex ante if she can determine that she should choose f at the interim stage anyway.

The ranking $\succeq_*$ in my decision framework has other possible interpretations. For example, GMMS define $\succeq_*$ in terms of "objective rationality" so that each preference $f \succeq_* g$ is "justified or defended on more or less objective grounds," and the decision maker can "convince others that she is right" when she chooses f rather than g. In this interpretation, the ranking $\succeq_*$ cannot be observed through choice behavior alone, but my representation results can still apply.

2.2 *Axioms and main representation result*

Consider the following axioms for the preferences $\succeq$ and $\succeq_*$.

Axiom 1 (Completeness). $\succeq$ *is complete.*

This condition requires that the decision maker can choose between any two acts f and g if she has no other alternatives available ex ante. Note that if she strictly prefers to postpone this decision, then her firm preference $\succeq_*$ is incomplete.

Axiom 2 (Consistency). *For all acts $f, g \in \mathscr{H}$, if $f \succeq_* g$, then $f \succeq g$.*

This axiom requires that if the decision maker chooses f rather than g when she can postpone this choice, then she should still prefer f to g when she has no option to wait.

Axiom 3 (Transitivity). $\succeq$ *and* $\succeq_*$ *are transitive.*

This condition requires that at the ex ante stage, the decision maker should neither have cyclical preferences nor expect that her interim preferences can be cyclical.

Say that $\succeq$ is *continuous* if the sets

$$\{\alpha \in [0,1] : \alpha f + (1-\alpha) g \succeq h\} \quad \text{and} \quad \{\alpha \in [0,1] : \alpha f + (1-\alpha) g \preceq h\}$$

are closed in $[0,1]$. Adopt the same notion of continuity for $\succeq_*$.

Axiom 4 (Continuity). *$\succeq$ and $\succeq_*$ are continuous.*

As customary, continuity is motivated by an abstract mathematical intuition that cannot be refuted by any finite number of observations.

Given acts $f, g \in \mathscr{H}$, write $f \geqq g$ if $f(s) \succeq g(s)$ for all $s \in S$.

Axiom 5 (Monotonicity). *For all acts $f, g \in \mathscr{H}$, if $f \geqq g$, then $f \succeq_* g$.*

Monotonicity assumes that the decision maker's *risk attitude*—that is, her ex ante ranking of lotteries in $\mathscr{L}$—should be preserved at the interim and ex post stages. Then the firm preference $f \succeq_* g$ is intuitive for all acts f and g such that $f \geqq g$. By Monotonicity and Consistency, the firm preference $\succeq_*$ coincides with $\succeq$ on the domain $\mathscr{L}$ of all constant acts.

Axiom 6 (Firm Independence). *For all $\alpha \in [0,1]$ and acts $f, g, h \in \mathscr{H}$,*

$$f \succeq_* g \quad \Rightarrow \quad \alpha f + (1-\alpha) h \succeq_* \alpha g + (1-\alpha) h.$$

To motivate this axiom, suppose that the decision maker expects to rank all acts via expected utility at the interim stage. Then she should expect her interim preference between any acts $f, g \in \mathscr{H}$ to be unaffected if f and g are both mixed with a common weight $\alpha \in [0,1]$ and a common act $h \in \mathscr{H}$.

By contrast, the ex ante preference $\succeq$ may still violate the expected utility model because of ambiguity aversion. Even in this case, the decision maker may still plan to comply with expected utility at the interim stage, providing that she expects to have sufficiently detailed information by that time. For example, expected utility maximization is plausible in Ellsberg's setting after the precise composition of the urn is announced.

Say that f is *more secure* than g (or g is *less secure* than f) if for all $l \in \mathscr{L}$,

$$g \succeq_* l \quad \Rightarrow \quad f \succeq_* l.$$

That is, if f is firmly preferred to any lottery l that g is firmly preferred to.

Being ambiguity averse, the decision maker may be biased in favor of more secure acts. This bias can be motivated in part by the *comparative ignorance* effect. Fox and Tversky (1995) observe that ambiguity aversion is much more common empirically when subjects compare "their limited knowledge about an event with their superior knowledge about another event." Similarly, the decision maker's ambiguity aversion in my framework can rely on the comparison between her limited ex ante knowledge about acts f and g and her superior knowledge about lotteries l.

Axiom 7 (Cautious Independence). *For all $\alpha \in [0,1]$, acts $f, g, h \in \mathcal{H}$ and lotteries $l \in \mathcal{L}$ such that $\alpha f + (1-\alpha)h$ is more secure than $\alpha f + (1-\alpha)l$, but $\alpha g + (1-\alpha)h$ is less secure than $\alpha g + (1-\alpha)l$,*

$$f \succeq g \quad \Rightarrow \quad \alpha f + (1-\alpha)h \succeq \alpha g + (1-\alpha)h. \tag{3}$$

This axiom implies both Certainty Independence and Uncertainty Aversion that GS formulate in their multiple priors model.[3] Indeed, if $h = l$, then by Cautious Independence,

$$f \succeq g \quad \Rightarrow \quad \alpha f + (1-\alpha)l \succeq \alpha g + (1-\alpha)l,$$

which is equivalent to Certainty Independence.

Moreover, for any $\alpha \in [0,1]$ and $f, g \in \mathcal{H}$, there exists $l \in \mathcal{L}$ such that $\alpha f + (1-\alpha)g$ is more secure than $\alpha f + (1-\alpha)l$, but g is less secure than $\alpha g + (1-\alpha)l$. (The existence of such lottery follows from Axioms 1–6 and is explained later.) Then Cautious Independence implies

$$f \succeq g \quad \Rightarrow \quad \alpha f + (1-\alpha)g \succeq g$$

and hence, Uncertainty Aversion.[4]

In addition, Cautious Independence requires that the decision maker should preserve her preference $\alpha f + (1-\alpha)l \succeq \alpha g + (1-\alpha)l$ when the constant act l in these mixtures is replaced by any act h such that $\alpha f + (1-\alpha)h$ is more secure than $\alpha f + (1-\alpha)l$, but $\alpha g + (1-\alpha)h$ is less secure than $\alpha g + (1-\alpha)l$. This separability is intuitive if the decision maker's ambiguity aversion is based exclusively on her bias in favor of more secure acts.

Note that both rankings $\succeq$ and $\succeq_*$ are essential for the statement of Cautious Independence. To clarify the role of the firm preference $\succeq_*$, let $u \in \mathcal{U}$ be an expected utility representation for the ranking of lotteries. For any act $f \in \mathcal{H}$, define its *security level*

$$U_*(f) = \max_{l' \preceq_* f} u(l')$$

as the maximal interim utility that the act f guarantees ex ante according to the firm preference $\succeq_*$. For any $\alpha \in [0,1]$ and acts $f, g \in \mathcal{H}$, define the *security premium* of the mixture $\alpha f + (1-\alpha)g$ by

$$SP(\alpha, f, g) = U_*(\alpha f + (1-\alpha)g) - [\alpha U_*(f) + (1-\alpha)U_*(g)],$$

which is roughly analogous to the definition of *risk premium* for preferences over monetary gambles (see Kreps 1988, p. 74). Given Axioms 1–6,

[3]More broadly, Certainty Independence is satisfied by *invariant biseparable preferences*, which are proposed by Ghirardato et al. (2004, 2005) to accommodate various attitudes towards ambiguity. On the other hand, Certainty Independence can be violated if $\succeq$ has a *variational* utility representation, as axiomatized by Maccheroni et al. (2006).

[4]In the multiple priors model, it is sufficient to impose Uncertainty Aversion only for symmetric mixtures with $\alpha = \frac{1}{2}$. Similarly, my representation results are unaffected if $\alpha = \frac{1}{2}$ is fixed in Cautious Independence.

- the functions u and U_* are well-defined
- f is more secure than g if and only if $U_*(f) \geq U_*(g)$
- for all $\alpha \in [0,1]$ and acts $f, g, h \in \mathscr{H}$,

$$SP(\alpha, f, h) \geq SP(\alpha, g, h)$$

  if and only if there exists $l \in \mathscr{L}$ such that $\alpha f + (1-\alpha)h$ is more secure than $\alpha f + (1-\alpha)l$, but $\alpha g + (1-\alpha)h$ is less secure than $\alpha g + (1-\alpha)l$. (See Lemma 1 in Appendix A.)

Therefore, Cautious Independence requires separability (3) whenever the security premium $SP(\alpha, f, h)$ is greater or equal than $SP(\alpha, g, h)$. In this case, one can interpret the mixture $\alpha f + (1-\alpha)h$ to be a *better hedge* than $\alpha g + (1-\alpha)h$.

Say that $\succeq$ is *extremely cautious* if for all $f \in \mathscr{H}$ and $l \in \mathscr{L}$,

$$f \succ l \quad \Rightarrow \quad f \succeq_* l.$$

This condition is the same as Caution in GMMS. It requires that any act f that is not firmly preferred to a constant act l should not be strictly preferred to l either.

The following theorem is my main representation result.

THEOREM 1. *Preferences $\succeq$ and $\succeq_*$ satisfy Axioms 1–7 if and only if there exist $\Delta \in \mathscr{C}$, $p \in \Delta$, $\varepsilon \in [0,1]$, and $u \in \mathscr{U}$ such that $\succeq$ is represented by*

$$U(f) = (1-\varepsilon)u\left(f(p)\right) + \varepsilon \min_{q \in \Delta} u\left(f(q)\right), \tag{4}$$

and for all acts $f, g \in \mathscr{H}$,

$$f \succeq_* g \quad \Leftrightarrow \quad u\left(f(q)\right) \geq u\left(g(q)\right) \quad \text{for all } q \in \Delta. \tag{5}$$

Moreover, if $\succeq$ is not extremely cautious, and $\succeq$ and $\succeq_*$ have representations (4) *and* (5) *with other components $u' \in \mathscr{U}$, $\Delta' \in \mathscr{C}$, $\varepsilon' \in [0,1]$, and $p' \in \Delta'$, then $\Delta' = \Delta$, $\varepsilon' = \varepsilon$, $p' = p$, and $u' = \alpha u + \beta$ for some $\alpha > 0$ and $\beta \in \mathbb{R}$.*

This theorem delivers a joint model for the pair of rankings $\succeq$ and $\succeq_*$. The Bewley-type representation (5) for the firm preference $\succeq_*$ asserts that $f \succeq_* g$ if and only if the expected utility of f is greater than or equal to the expected utility of g for all probabilistic beliefs in the set Δ. This set can be interpreted as the collection of all interim beliefs that the decision maker views as possible ex ante. Formally, Δ is derived from the firm preference $\succeq_*$ via GMM's results.

The main novelty of Theorem 1 is the utility representation (4) for the preference $\succeq$. This representation evaluates every act f via the ε-mixture of the least favorable belief in the set Δ and a probability measure $p \in \Delta$ that is common for all acts f. Then it is intuitive to interpret p as the decision maker's ex ante probabilistic belief, and the weight ε as a degree of her aversion towards ambiguity of her interim beliefs in the set

Δ. Alternatively, $1-\varepsilon$ can be viewed as an index of ex ante subjective confidence in the belief p.

In particular, if $\varepsilon=0$, then $\succeq$ is represented by expected utility $u\left(f(p)\right)$. In this case, the decision maker exhibits no ambiguity aversion at the ex ante stage, even though her firm preference $\succeq_*$ may be incomplete, and the corresponding set Δ may be non-singleton.

By contrast, if $\varepsilon=1$, then $\succeq$ is extremely cautious, and it is represented by maxmin expected utility

$$U(f)=\min_{q\in\Delta} u\left(f(q)\right),$$

which does not depend on p at all. Theorem 1 implies that this representation together with (5) is equivalent to the combination of Axioms 1–6, Certainty Independence, and extreme caution.[5] Therefore, the main representation result (Theorem 3) in GMMS can be obtained as a corollary of Theorem 1.

In general, representation (4) has the maxmin expected utility form

$$U(f)=\min_{q\in\Pi} u\left(f(q)\right), \tag{6}$$

where the set

$$\Pi=(1-\varepsilon)\{p\}+\varepsilon\Delta \tag{7}$$

is the ε*-contamination* of the probability measure $p\in\Delta$ by the convex and closed set Δ. Besides the added structure (7) for the set Π, Theorem 1 differs from the multiple priors model in several respects.

First, my framework requires the firm preference $\succeq_*$ as an extra primitive. If $\succeq_*$ is not given, then any convex and closed set Π can be written in the form (7), most trivially for $\varepsilon=1$ and $\Delta=\Pi$. Therefore, for any preference $\succeq$ that has a maxmin expected utility representation, there exists a ranking $\succeq_*$ such that the pair $(\succeq,\succeq_*)$ complies with all the assumptions of Theorem 1.

Second, Cautious Independence implies both Certainty Independence and Uncertainty Aversion, but unlike these axioms, it cannot be stated in terms of $\succeq$ alone. For example, if $\succeq_*$ is complete, then $\Delta=\{p\}$, and $\succeq_*$ is represented by expected utility $u(f(p))$. In this case, Cautious Independence turns into standard Independence and requires separability (3) for all $\alpha\in[0,1]$ and $f,g,h\in\mathscr{H}$ because $SP(\alpha,f,h)=SP(\alpha,g,h)=0$. On

[5]To show this claim, suppose that $\succeq$ and $\succeq_*$ satisfy Axioms 1–6, Certainty Independence, and extreme caution. Then for all $f,g\in\mathscr{H}$ and $l\in\mathscr{L}$, if f is more secure than g, then

$$g\succ l \quad\Rightarrow\quad g\succeq_* l \quad\Rightarrow\quad f\succeq_* l \quad\Rightarrow\quad f\succeq l,$$

and hence, $f\succeq g$ (otherwise, $g\succ l\succ f$ for some $l\in\mathscr{L}$). Therefore, for all $f,g,h\in\mathscr{H}$ and $l\in\mathscr{L}$, if $\alpha f+(1-\alpha)h$ is more secure than $\alpha f+(1-\alpha)l$, but $\alpha g+(1-\alpha)h$ is less secure than $\alpha g+(1-\alpha)l$, then

$$f\succeq g \quad\Rightarrow\quad \alpha f+(1-\alpha)h\succeq \alpha f+(1-\alpha)l\succeq \alpha g+(1-\alpha)l\succeq \alpha g+(1-\alpha)h.$$

Thus, $\succeq$ satisfies Cautious Independence. By Theorem 1, $\succeq$ and $\succeq_*$ have representations (4) and (5). It is easy to check that if Δ is not a singleton and $\varepsilon<1$, then $\succeq$ is not extremely cautious. On the other hand, if Δ is a singleton or $\varepsilon=1$, then $\succeq$ has the required maxmin expected utility representation.

the other hand, any preference reversal $f \succeq g$ and $\alpha f + (1-\alpha)h \prec \alpha g + (1-\alpha)h$ that is consistent with the multiple priors model is also consistent with Cautious Independence for some ranking $\succeq_*$.

Third, my model interprets the set Δ rather than Π as the collection of all probabilistic beliefs that the decision maker deems possible ex ante. Therefore, Theorem 1 pins down a "subjective state space" Δ that cannot be properly identified within the general multiple priors model. Note though that representations (4) and (5) are unchanged if Δ is replaced by any other closed set $\Delta_* \neq \Delta$ as long as the convex hull of Δ_* coincides with Δ. Therefore, my model does not distinguish between two individuals who have the same u, p, and ε, and whose interim beliefs vary in different sets Δ and Δ_* with common convex hulls.

2.3 *Sketch of proof*

Next, I sketch the construction of representation (4) for the preference $\succeq$. This construction is the hardest part of the proof of Theorem 1.

Suppose that $\succeq$ and $\succeq_*$ satisfy Axioms 1–7.

Step 1. By Herstein and Milnor's (1953) Theorem, the ranking of lotteries has an expected utility representation $u \in \mathscr{U}$ that is unique up to a positive linear transformation. Without loss of generality, assume that u is non-constant. By Ghirardato et al. (2004, Proposition A.2) there is a unique set $\Delta \in \mathscr{C}$ such that $\succeq_*$ has the required representation (5). Note that for all $f \in \mathscr{H}$, $U_*(f) = \min_{q\in\Delta} u(f(q))$.

Step 2. Show that $\succeq$ satisfies all the conditions in Theorem 1 in GS. The concavity of the function U_* implies that for all $\alpha \in [0,1]$ and $f, g \in \mathscr{H}$,

$$SP(\alpha, f, g) \geq 0 = SP(\alpha, g, g).$$

Then there exists $l \in \mathscr{L}$ such that $\alpha f + (1-\alpha)h$ is more secure than $\alpha f + (1-\alpha)l$, but $\alpha g + (1-\alpha)h$ is less secure than $\alpha g + (1-\alpha)l$. (See Lemma 1 in Appendix A.) By Cautious Independence, $\succeq$ satisfies Uncertainty Aversion. The other conditions in GS's list are obvious. Thus, there is a unique set $\Pi \in \mathscr{C}$ such that $\succeq$ has the maxmin expected utility representation (6).

Step 3. Assume that S is finite and $\Sigma = 2^S$ (the general case is treated separately). Let $a \cdot b$ denote the scalar product of the vectors $a, b \in \mathbb{R}^S$. For every vector $a \in \mathbb{R}^S$, let

$$V(a) = \min_{q\in\Delta} q \cdot a \quad \text{and} \quad W(a) = \min_{q\in\Pi} q \cdot a.$$

Let $\mathbb{D}$ be the set of all points $a \in \mathbb{R}^S$ where the functions V and W are both differentiable. For all $a \in \mathbb{D}$, let

$$v(a) = \nabla V(a) \quad \text{and} \quad w(a) = \nabla W(a).$$

Then for all $a \in \mathbb{D}$,

$$V(a) = v(a) \cdot a \quad \text{and} \quad W(a) = w(a) \cdot a. \tag{8}$$

Moreover, for any $a, b \in \mathbb{D}$, there is $\varepsilon \geq 0$ such that

$$w(a) - w(b) = \varepsilon(v(a) - v(b)).$$

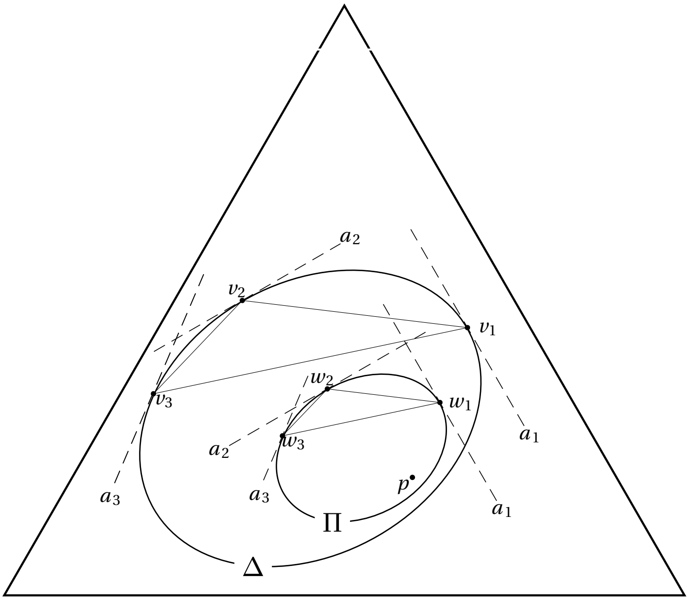

FIGURE 2. The construction of ε and p in Theorem 1.

This claim is derived from Cautious Independence (see Lemma 2 below). Roughly, the argument proceeds as follows. Suppose that $v(a)-v(b)$ is not parallel to $w(a)-w(b)$. Note that if $a-a'$ is a constant vector, then $v(a)=v(a')$ and $w(a)=w(a')$. Therefore, it is without loss of generality to assume that $W(a)=W(b)$ and $V(c)=0$. By the separating hyperplane theorem, there is a vector c such that

$$(v(a)-v(b))\cdot c>0>(w(a)-w(b))\cdot c.$$

If α is sufficiently small, then

$$\begin{aligned}V(\alpha c+(1-\alpha)a)-(\alpha V(c)+(1-\alpha)V(a))&=\alpha v(a)\cdot c+o(\alpha)\\&>\alpha v(b)\cdot c+o(\alpha)=V(\alpha c+(1-\alpha)b)-(\alpha V(c)+(1-\alpha)V(b)).\end{aligned}$$

Similarly, $W(\alpha c+(1-\alpha)a)<W(\alpha c+(1-\alpha)b)$ because $W(a)=W(b)$. The former inequality implies that $\alpha c+(1-\alpha)a$ is a "better hedge" than $\alpha c+(1-\alpha)b$ according to the security level function V. Yet the latter inequality is a reversal of $W(a)\geq W(b)$. This situation translates into a contradiction with Cautious Independence.

Step 4. Take any $a_1,a_2,a_3\in\mathbb{D}$ and let $w_i=w(a_i)$ and $v_i=v(a_i)$ for $i=1\ldots3$. By Step 3, the triangles $v_1v_2v_3$ and $w_1w_2w_3$ are homothetic because their edges v_iv_j and w_iw_j are parallel. Figure 2 illustrates this geometric intuition for a three-element state space. Therefore, there is $\varepsilon\geq 0$ and a vector $p\in\mathbb{R}^S$ such that for all $a\in\mathbb{D}$,

$$w(a)=\varepsilon v(a)+(1-\varepsilon)p.$$

By (8), $W(a) = \varepsilon V(a) + (1-\varepsilon) p \cdot a$ for all $a \in \mathbb{D}$. By Rockafellar (1970, Theorem 25.5), $\mathbb{D}$ is dense in all of $\mathbb{R}^S$. By continuity,

$$W(a) = \varepsilon V(a) + (1-\varepsilon) p \cdot a$$

for all $a \in \mathbb{R}^S$. Thus, $\Pi = \varepsilon \Delta + (1-\varepsilon) p$. By Consistency, $W(a) \geq V(a)$ for all $a \in \mathbb{R}^S$. It follows that $\varepsilon \leq 1$ and $p \in \Delta$. (See Lemma 3 below.)

Step 5. Extend the utility representation (4) to an arbitrary state space (S, Σ) by extending the probability measure p from finite subalgebras of Σ to the entire Σ.

2.4 *Behavioral meaning and elicitation of p and ε*

Theorem 1 delivers a unique probability measure p as a formal component of the utility representation for the preference $\succeq$. In contrast with the standard models of subjective probability, p is not revealed by the decision maker's betting preference. The behavioral meaning of the belief p is clarified by the following result.

Theorem 2. *Suppose that $\succeq$ and $\succeq_*$ satisfy Axioms 1–7, and $\succeq$ is not extremely cautious. Then for all $f, g \in \mathscr{H}$ such that f is more secure than g,*

$$f(p) \succ g(p) \quad \Rightarrow \quad f \succ g, \tag{9}$$

and p is the only probability measure in $\mathscr{P}$ that satisfies this condition.

Therefore, the subjective probabilistic belief p manifests itself directly through the ex ante rankings (9) and is uniquely determined by these rankings (except for the case of extreme caution). More precisely, p is the only probability measure in the entire simplex $\mathscr{P}$ such that for all acts f and g, if f is more secure than g, and if $f(p) \succ g(p)$, then $f \succ g$. In contrast to the expected utility theory or Machina and Schmeidler's (1992) probabilistic sophistication, representation (4) does not imply condition (9) for all acts f and g. Even if $f(p) \succ g(p)$, the decision maker may still strictly prefer g to f when g is more secure than f. Intuitively, this preference reversal occurs because she is not fully confident of her subjective assessment of probabilities p and hence is biased in favor of more secure acts.

To interpret the parameter ε in representations (5) and (6), compare two pairs of rankings $(\succeq, \succeq_*)$ and $(\succeq', \succeq_*)$ that share the same firm preference. Say that $\succeq'$ is *more averse to interim ambiguity* than $\succeq$ if for all acts $f, g \in \mathscr{H}$ such that f is more secure than g,

$$f \succeq g \quad \Rightarrow \quad f \succeq' g. \tag{10}$$

Put differently, this condition requires that $\succeq$ may strictly prefer an act g to a more secure act f only if $\succeq'$ does so. It follows that for all acts $f \in \mathscr{H}$ and lotteries $l \in \mathscr{L}$,

$$l \succeq g \quad \Rightarrow \quad l \succeq' g$$

because the ranking $l \succeq g$ implies that the lottery l is more secure than the act g. Thus, condition (10) strengthens the comparative definition of ambiguity aversion in Ghirardato and Marinacci (2002).

Theorem 3. *Suppose that both* $(\succeq, \succeq_*)$ *and* $(\succeq', \succeq_*)$ *satisfy Axioms 1–7. Then* $\succeq'$ *is more averse to interim ambiguity than* $\succeq$ *if and only if* $\succeq$ *and* $\succeq'$ *have representations* (4) *with tuples* $(u, \Delta, p, \varepsilon)$ *and* $(u, \Delta, p, \varepsilon')$ *such that* $\varepsilon' \geq \varepsilon$.

This result describes the behavioral effect of a ceteris paribus increase in the parameter ε in my model.

The above geometric construction and behavioral interpretations of p and ε do not provide any practical methods to derive these components from empirical data. To elicit p and ε from the preferences $\succeq$ and $\succeq_*$, one may proceed as follows.

Take any payoffs $x \succ y$, such as $x = \$100$ and $y = \$0$. For any $\gamma \in [0,1]$, let l_γ be a lottery that delivers x and y with probabilities γ and $1-\gamma$ respectively. For any event $A \in \Sigma$, let

$$\pi(A) = \max\{\gamma \in [0,1] : x\,A\,y \succeq l_\gamma\}$$
$$\pi_*(A) = \max\{\gamma \in [0,1] : x\,A\,y \succeq_* l_\gamma\}.$$

Here the values $\pi(A)$ and $\pi_*(A)$ measure the decision maker's ex ante willingness to bet on the event A in two different situations: in the former case, she must choose between the bet $x\,A\,y$ and the lottery l_γ immediately, and in the latter, she can postpone this choice until the interim stage.

By Consistency, $\pi_*(A) \leq \pi(A)$. Consider three possible cases.

(i) For all $A \in \mathscr{A}$, $\pi(A) + \pi(A^c) = 1$. Then $\succeq$ is represented by expected utility with $p = \pi$, and $\varepsilon \in [0,1]$ is arbitrary.

(ii) There is $A \in \mathscr{A}$ such that $\pi_*(A) + \pi_*(A^c) = \pi(A) + \pi(A^c) < 1$. Take $\varepsilon = 1$ and arbitrary $p \in \Delta$. In this case, $\succeq$ is extremely cautious.

(iii) There is $A \in \mathscr{A}$ such that $\pi_*(A) + \pi_*(A^c) < \pi(A) + \pi(A^c) < 1$. By (7),

$$\pi(A) = (1-\varepsilon)p(A) + \varepsilon\pi_*(A)$$
$$\pi(A^c) = (1-\varepsilon)(1-p(A)) + \varepsilon\pi_*(A^c).$$

By summing the two equations, we obtain

$$\varepsilon = \frac{1-\pi(A)-\pi(A^c)}{1-\pi_*(A)-\pi_*(A^c)}.$$

For all events $B \in \mathscr{A}$, take

$$p(B) = \frac{\pi(B) - \varepsilon\pi_*(B)}{1-\varepsilon}.$$

In this case, $\succeq$ is not extremely cautious, and both $\varepsilon < 1$ and p are determined uniquely.

Note that the above construction of ε and p can be sensitive to small measurement errors in the capacities π and π_* when the ε is close to 1, or the capacity π is close to the capacity π_*. The robustness of ε and p is discussed in Appendix B.

2.5 *Complete ignorance*

An important special case of Theorem 1 is obtained when the firm preference $\succsim_*$ satisfies

$$f \succsim_* g \quad \Leftrightarrow \quad f(s) \succsim g(s) \quad \text{for all } s \in S \tag{11}$$

for all acts $f, g \in \mathscr{H}$. This equivalence is intuitive if the decision maker is *completely ignorant* ex ante about her interim beliefs. For example, she may have no ex ante information about the composition of Ellsberg's urn, but expect to learn this composition precisely at the interim stage. Equivalence (11) is also intuitive if the decision maker views all states in S as possible ex ante, but expects to learn the true state at the interim stage.

Note that if (11) is assumed, then the firm preference $\succsim_*$ need not be taken as a formal primitive in Theorem 1. In this case, an act f is more secure than g if and only if there is $s \in S$ such that $f(s') \succsim g(s)$ for all $s' \in S$. Then Theorem 1 asserts that the preference $\succsim$ is complete, transitive, continuous, monotonic, and satisfies Cautious Independence if and only if $\succsim$ has the utility representation

$$U(f) = (1-\varepsilon)u(f(p)) + \varepsilon \min_{s \in S} u(f(s)) \tag{12}$$

where $\varepsilon \in [0,1]$, $p \in \mathscr{P}$, and $u \in \mathscr{U}$. This function has the maxmin expected utility form with the set of priors

$$\Pi = (1-\varepsilon)\{p\} + \varepsilon \mathscr{P}. \tag{13}$$

This model has been used in economic applications to asset pricing (Epstein and Wang 1994), insurance (Carlier et al. 2003), and search (Nishimura and Ozaki 2004). If fact, some of these authors use the term ε*-contamination* to refer to the structure (13) rather than to the general (7).

Representation (12) has other noteworthy distinctions. First, if $\succsim$ is not extremely cautious, then for all lotteries $l \succ l'$ and all events $A \subsetneq S$ and $B \subsetneq S$,

$$p(A) > p(B) \quad \Leftrightarrow \quad l\,A\,l' \succ l\,B\,l'.$$

Therefore, the probability measure p represents the comparative likelihood relation induced by the preference $\succsim$ over all events $A \subsetneq S$. By contrast, the comparative likelihood relation does not have an additive representation in the Ellsberg Paradox.

Second, representation (12) has the *Choquet expected utility* form

$$U(f) = \int_S u(f(s))\, d\pi,$$

where the capacity $\pi : \Sigma \to [0,1]$ is such that $\pi(S) = 1$ and $\pi(A) = (1-\varepsilon)p(A)$ for all events $A \subsetneq S$. By contrast, if the set Π does not equal the core of some convex capacity, then representation (4) does not comply with Choquet expected utility.

3. Possible extensions

As mentioned above, representations (4) and (5) remain intuitive as long as the decision maker expects her interim beliefs to vary in any set Δ_* that has the convex hull Δ. To identify the true subjective state space Δ_*, one may consider a preference $\succeq_0$ over all *menus* of acts (i.e. subsets of $\mathscr{H}$). Note that both $\succeq$ and $\succeq_*$ can be derived from $\succeq_0$ by taking

$$\begin{aligned} f \succeq g \quad &\Leftrightarrow \quad \{f\} \succeq_0 \{g\}, \\ f \succeq_* g \quad &\Leftrightarrow \quad \{f\} \succeq_0 \{f, g\} \end{aligned}$$

for all acts $f, g \in \mathscr{H}$. The rankings $\{f\} \succeq_0 \{g\}$ and $\{f\} \succeq_0 \{f, g\}$ mean respectively that (i) the decision maker prefers to commit to f rather than to g ex ante, and (ii) she is willing to commit to f even if she can keep both f and g feasible until the interim stage.

However, there is not a unique way to extend my model to preferences over menus. To illustrate, adopt the approach of Epstein et al. (2008). Let $\varepsilon < 1$, let μ be a probability measure on $\mathscr{P}$ with support Δ_μ, and let the preference $\succeq_0$ over all closed[6] menus $A \subset \mathscr{H}$ be represented by

$$U_0(A) = (1-\varepsilon)\int_{\mathscr{P}} \max_{f\in A} u(f(q))\, d\mu + \varepsilon \min_{q\in\Delta_\mu} \max_{f\in A} u(f(q)).$$

The decision maker as portrayed by this representation expects that her interim preference will comply with expected utility and her interim belief will belong to the set Δ_μ. She aggregates these beliefs via a second-order belief μ, but being ambiguity averse, she puts an additional weight ε on the belief $q \in \Delta_\mu$ that is least favorable for the menu at hand. Consider another preference $\succeq_0'$ that is represented by

$$U_0'(A) = (1-\varepsilon)\int_{\mathscr{P}} \max_{f\in A} u(f(q))\, d\mu + \varepsilon \max_{f\in A} \min_{q\in\Delta_\mu} u(f(q)).$$

Here the decision maker behaves as if she expects that with probability $1-\varepsilon$, her interim preference will comply with expected utility and her interim belief will be resolved in the set Δ_μ according to the second-order distribution μ. Yet she also expects that with probability ε, she will get no additional information at the interim stage and will have maxmin expected utility $\min_{q\in\Delta_\mu} u(f(q))$. In this case, her "subjective states" include some rankings that violate expected utility altogether.

The utility functions U_0 and U_0' correspond to the models of short-run coarseness and persistent coarseness in Epstein et al. (2008). Besides the difference in interpretations, these models have distinct behavioral properties. For example, the preference $\succeq_0$ satisfies Indifference to Randomization of menus, but $\succeq_0'$ does not.

Yet both $\succeq_0$ and $\succeq_0'$ induce the same pair of rankings $\succeq$ and $\succeq_*$ that have representations (4) and (5) with $p = \int_{\mathscr{P}} q\, d\mu$ and Δ equal to the convex hull of Δ_μ. Indeed, for all $f \in \mathscr{H}$,

$$U(\{f\}) = U'(\{f\}) = (1-\varepsilon)u(f(p)) + \varepsilon \min_{q\in\Delta_\mu} u(f(q))$$

[6]For simplicity, assume that S and X are finite, and the spaces $\mathscr{L}$ and $\mathscr{H}$ have Euclidean topologies.

and hence $\succeq_0$ and $\succeq_0'$ induce the same $\succeq$. Moreover,

- if $u(f(q)) \geq u(g(q))$ for all $q \in \Delta_\mu$, then
$$U(\{f,g\}) = U(\{f\}) = U'(\{f\}) = U'(\{f,g\})$$
- if $u(g(q)) > u(f(q))$ for some $q \in \Delta_\mu$, then $u(g(q')) > u(f(q'))$ for all q' in some neighborhood N_q of q such that $\mu(N_q) > 0$ and hence $U(\{f,g\}) > U(\{f\})$ and $U'(\{f,g\}) > U'(\{f\})$.

Therefore, $\succeq_0$ and $\succeq_0'$ induce the same firm preference $\succeq_*$.

This example illustrates that my model has various possible extensions to preferences over menus, and these extensions provide different specifications for subjective states. The compensation for this lack of sharpness is the weaker primitives and weaker behavioral assumptions that are required to work with a pair of rankings $\succeq$ and $\succeq_*$ of individual acts rather than with a preference $\succeq_0$ over all menus of acts.

Alternatively, one may extend my representation results for different interpretations of the set Δ. In a companion paper (Kopylov 2008), I define Δ in terms of observable prices at which uncertain prospects can be sold in incomplete markets. An interesting distinction of this approach is that the decision maker need not agree with the market pricing, and her subjective belief p need not belong to Δ.

Finally, one may try to accommodate different attitudes towards ambiguity. It is straightforward to flip my model so that the preference $\succeq$ is represented by

$$U(f) = (1-\gamma)u\left(f(p)\right) + \gamma \max_{q\in\Delta} u\left(f(q)\right), \tag{14}$$

where $p \in \Delta$, $\Delta \in \mathscr{C}$, $u \in \mathscr{U}$, and $\gamma \in [0,1]$. To do so, say that an act f has *more potential* than g (or equivalently, g has *less potential* than f) if for all $l \in \mathscr{L}$,

$$l \succeq_* f \quad \Rightarrow \quad l \succeq_* g.$$

Rewrite Cautious Independence and impose invariance (3) for all $\alpha \in [0,1]$, acts f, g, h and lotteries l such that $\alpha f + (1-\alpha)h$ has more potential than $\alpha f + (1-\alpha)l$, but $\alpha g + (1-\alpha)h$ has less potential than $\alpha g + (1-\alpha)l$. Then Theorem 1 turns into a characterization of (14).

To accommodate a mixture of pessimistic and optimistic attitudes towards ambiguity, one can use a utility representation

$$U(f) = (1-\varepsilon-\gamma)u\left(f(p)\right) + \varepsilon \min_{q\in\Delta} u\left(f(q)\right) + \gamma \max_{q\in\Delta} u\left(f(q)\right),$$

where $\varepsilon, \gamma \in [0,1]$ are such that $\varepsilon + \gamma \leq 1$. A special case of this model for $\Delta = \mathscr{P}$ has been characterized by Chateauneuf et al. (2007), but the general case is an open research problem.

Appendices

A. Proofs of theorems 1–3

Suppose that $\succeq_*$ is represented by (5) for some $u \in \mathscr{U}$ and $\Delta \in \mathscr{C}$. For all $f \in \mathscr{H}$, let $U_*(f) = \max_{l : l \preceq_* f} u(l)$.

Lemma 1. *For all $\alpha \in [0,1]$ and $f, g, h \in \mathscr{H}$,*

(i) $U_*(f) = \min_{q \in \Delta} u(f(q))$

(ii) f is more secure than g if and only if $U_*(f) \geq U_*(g)$

(iii) $SP(\alpha, f, h) \geq SP(\alpha, g, h)$ if and only if there is $l \in \mathscr{L}$ such that $\alpha f + (1-\alpha)h$ is more secure than $\alpha f + (1-\alpha)l$ and $\alpha g + (1-\alpha)h$ is less secure than $\alpha g + (1-\alpha)l$.

Proof. The first two claims are obvious. Turn to (iii). By definition of security premia, for all $\alpha \in [0,1]$, $f, g, h \in \mathscr{H}$, and $l \in \mathscr{L}$,

$$\begin{aligned} SP(\alpha, f, h) - SP(\alpha, g, h) = {} & [U_*\left(\alpha f + (1-\alpha)h\right) - U_*\left(\alpha f + (1-\alpha)l\right)] \\ & + [U_*\left(\alpha g + (1-\alpha)l\right) - U_*\left(\alpha g + (1-\alpha)h\right)]. \end{aligned}$$

Therefore, if $\alpha f + (1-\alpha)h$ is more secure than $\alpha f + (1-\alpha)l$ and $\alpha g + (1-\alpha)h$ is less secure than $\alpha g + (1-\alpha)l$, then $SP(\alpha, f, h) \geq SP(\alpha, g, h)$. Conversely, suppose that $SP(\alpha, f, h) \geq SP(\alpha, g, h)$. Take $l \in \mathscr{L}$ such that $U_*\left(\alpha f + (1-\alpha)h\right) = U_*\left(\alpha f + (1-\alpha)l\right)$. Then $\alpha f + (1-\alpha)h$ is more secure than $\alpha f + (1-\alpha)l$ and

$$U_*\left(\alpha g + (1-\alpha)l\right) - U_*\left(\alpha g + (1-\alpha)h\right) = SP(\alpha, f, h) - SP(\alpha, g, h) \geq 0.$$

That is, $\alpha g + (1-\alpha)h$ is less secure than $\alpha g + (1-\alpha)l$. □

Suppose next that $\succeq$ is represented by

$$U(f) = (1-\varepsilon)u(f(p)) + \varepsilon U_*(f)$$

for some $\varepsilon \in [0,1]$ and $p \in \Delta$. Then Axioms 1–6 are easy to check. To verify Cautious Independence, take $\alpha \in [0,1]$, acts $f, g, h \in \mathscr{H}$, and a lottery $l \in \mathscr{L}$ such that $f \succeq g$, $\alpha f + (1-\alpha)h$ is more secure than $\alpha f + (1-\alpha)l$, but $\alpha g + (1-\alpha)h$ is less secure than $\alpha g + (1-\alpha)l$. Then

$$\begin{aligned} d_1 &= U\left(\alpha f + (1-\alpha)l\right) - U\left(\alpha g + (1-\alpha)l\right) \geq 0 \\ d_2 &= U_*\left(\alpha f + (1-\alpha)h\right) - U_*\left(\alpha f + (1-\alpha)l\right) \geq 0 \\ d_3 &= U_*\left(\alpha g + (1-\alpha)l\right) - U_*\left(\alpha g + (1-\alpha)h\right) \geq 0 \\ U\left(\alpha f + (1-\alpha)h\right) - U\left(\alpha g + (1-\alpha)h\right) &= d_1 + \varepsilon(d_2 + d_3) \geq 0 \end{aligned}$$

and hence $\alpha f + (1-\alpha)h \succeq \alpha g + (1-\alpha)h$.

Turn to sufficiency. Suppose that preferences $\succeq$ and $\succeq_*$ satisfy Axioms 1–7. By Herstein and Milnor's Theorem, the ranking of lotteries has an expected utility representation $u \in \mathscr{U}$ that is unique up to a positive linear transformation. If u is constant, then (4) and (5) are trivial. Hereafter, assume that u is non-constant. Without loss of generality, the range of u contains the interval $[-1, 1]$. For any act $f \in \mathscr{H}$, let $u(f) \in \mathbb{R}^S$ be the composition of u and f.

By Ghirardato et al. (2004, Proposition A.2) and Gilboa et al. (2008, Theorem 1), there is a unique $\Delta \in \mathscr{C}$ such that $\succeq_*$ is represented by (5). The corresponding function U_* satisfies Lemma 1.

The preference $\succeq$ satisfies all the conditions in Theorem 1 in GS. In particular, for all $\alpha \in [0, 1]$ and $f, g \in \mathscr{H}$,

$$SP(\alpha, f, g) \geq 0 = SP(\alpha, g, g)$$

because U_* is concave. By Lemma 1 and Cautious Independence, $\succeq$ satisfies Uncertainty Aversion. Thus, there is a unique set $\Pi \in \mathscr{C}$ such that $\succeq$ is represented by

$$U(f) = \min_{q \in \Pi} u(f(q)).$$

Assume that S is finite and $\Sigma = 2^S$ (the general case is treated separately). For any $a \in \mathbb{R}^S$, let

$$V(a) = \min_{q \in \Delta} q \cdot a \quad \text{and} \quad W(a) = \min_{q \in \Pi} q \cdot a. \tag{15}$$

For any $\gamma \in \mathbb{R}$, let $\vec{\gamma} = (\gamma, \ldots, \gamma) \in \mathbb{R}^S$. Then the functions $V, W : \mathbb{R}^S \to \mathbb{R}$ are continuous and concave, and satisfy

$$V\left(\alpha a + \vec{\gamma}\right) = \alpha V(a) + \gamma \quad \text{and} \quad W\left(\alpha a + \vec{\gamma}\right) = \alpha W(a) + \gamma \tag{16}$$

for all vectors $a \in \mathbb{R}^S$ and scalars $\alpha \geq 0$, $\gamma \in \mathbb{R}$.

Next, I claim that for all $a \in \mathbb{R}^S$,

$$W(a) \geq V(a). \tag{17}$$

To show this claim, suppose that $W(a) < V(a)$ for some $a \in \mathbb{R}^S$. Without loss of generality, $a = u(f)$ for some $f \in \mathscr{H}$. Take $l \in \mathscr{L}$ such that $U_*(f) = V(a) > u(l) > W(f) = U(f)$. Then $f \succeq_* l$, but $l \succ f$, which contradicts Consistency.

Let $\mathbb{D}$ be the set of all points $a \in \mathbb{R}^S$ where the functions V and W are both differentiable. For every $a \in \mathbb{D}$, let

$$v(a) = \nabla V(a) \quad \text{and} \quad w(a) = \nabla W(a).$$

Take any $q_a \in \Delta$ such that $V(a) = q_a \cdot a$. Then $q_a = v(a)$ because for all $b \in \mathbb{R}^S$ and $\delta \in \mathbb{R}$,

$$V(a) + \delta(q_a \cdot b) = q_a \cdot (a + \delta b) \geq \min_{q \in \Delta} q \cdot (a + \delta b) = V(a + \delta b) = V(a) + \delta(v(a) \cdot b) + o(\delta),$$

and hence $q_a \cdot b = v(a) \cdot b$. Therefore, the vector $v(a) \in \Delta$ is the unique minimizer in (15): for all $q \in \Delta$ such that $q \neq v(a)$,

$$V(a) = v(a) \cdot a < q \cdot a. \tag{18}$$

Similarly, the vector $w(a)\in\Pi$ is the unique minimizer in (15): for all $q\in\Pi$ such that $q\neq w(a)$,

$$W(a)=w(a)\cdot a<q\cdot a. \tag{19}$$

It follows from (18) and (19) that $v(a)$ and $w(a)$ are extreme points in Δ and Π respectively.

Lemma 2. *For any $a,b\in\mathbb{D}$, there exists $\varepsilon\geq 0$ such that*

$$w(a)-w(b)=\varepsilon(v(a)-v(b)). \tag{20}$$

Proof. I claim that for all $a,b,c\in\mathbb{R}^S$ such that $V(a+c)\geq V(a)$ and $V(b+c)\leq V(b)$,

$$W(a)\geq W(b) \quad\Rightarrow\quad W(a+c)\geq W(b+c). \tag{21}$$

By (16), it is sufficient to show this claim for vectors $a,b,c\in[-1,1]^S$. Take acts $f,g,h\in\mathscr{H}$ such that $u(f)=a$, $u(g)=b$, and $u(h)=c$. Take a lottery $l\in\mathscr{L}$ such that $u(l)=0$. Then the inequalities $W(a)\geq W(b)$, $V(a+c)\geq V(a)$, and $V(b+c)\leq V(b)$ imply respectively that $f\succeq g$,

$$U_*\left(\tfrac{f+h}{2}\right)=V\left(u\left(\tfrac{f+h}{2}\right)\right)=V\left(\tfrac{a+c}{2}\right)\geq V\left(\tfrac{a}{2}\right)=V\left(u\left(\tfrac{f+l}{2}\right)\right)=U_*\left(\tfrac{f+h}{2}\right)$$
$$U_*\left(\tfrac{g+l}{2}\right)=V\left(u\left(\tfrac{g+l}{2}\right)\right)=V\left(\tfrac{b}{2}\right)\geq V\left(\tfrac{b+c}{2}\right)=V\left(u\left(\tfrac{g+h}{2}\right)\right)=U_*\left(\tfrac{g+h}{2}\right).$$

By Lemma 1, $\frac{1}{2}(f+h)$ is more secure than $\frac{1}{2}(f+l)$, but $\frac{1}{2}(g+l)$ is less secure than $\frac{1}{2}(g+h)$. By Cautious Independence, $\frac{1}{2}(f+h)\succeq\frac{1}{2}(g+h)$. Therefore

$$W\left(u\left(\tfrac{f+h}{2}\right)\right)\geq W\left(u\left(\tfrac{g+h}{2}\right)\right),$$

and by (16), $W(a+c)\geq W(b+c)$.

Turn to (20). Fix any $a,b\in\mathbb{D}$. The derivatives of the functions W and V, and hence the equality (20), are unaffected if the vectors a and b are replaced by $a-V(a)\vec{1}$ and $b-(W(b)+V(a)-W(a))\vec{1}$ respectively. Without loss of generality assume that $W(a)=W(b)$ and $V(a)=0$.

By the separation theorem, the convex hull of the vectors $v(a)$, $-v(b)$, and $w(b)-w(a)$ either contains 0 or can be separated from 0 by a hyperplane. Therefore, one of the following two cases must hold.

Case 1. There are $\lambda_1,\lambda_2,\lambda_3\geq 0$ such that $\lambda_1+\lambda_2+\lambda_3=1$ and

$$\lambda_1 v(a)-\lambda_2 v(b)+\lambda_3(w(b)-w(a))=0.$$

Then $\lambda_1=\lambda_2$ because $v(a)\cdot\vec{1}=v(b)\cdot\vec{1}=w(a)\cdot\vec{1}=w(b)\cdot\vec{1}=1$. If $\lambda_3\neq 0$, then (20) holds for $\varepsilon=\lambda_1/\lambda_3$. Suppose that $\lambda_3=0$. Then $\lambda_1=\lambda_2\neq 0$ and $v(a)=v(b)$. Recall that $V(a)=0$. Then for any $\delta>0$, $V(a+\delta a)=V(a)$ and

$$V(b+\delta a)=\min_{q\in\Delta} q\cdot(b+\delta a)\leq v(b)\cdot(b+\delta a)=V(b)$$

because $v(b)\cdot b = V(b)$ and $v(b)\cdot a = v(a)\cdot a = V(a) = 0$. Let $c = \delta a$. Then by (21), $W(a+\delta a) \geq W(b+\delta a)$. That is,

$$W(a)+\delta(w(a)\cdot a)+o(\delta) \geq W(b)+\delta(w(b)\cdot a)+o(\delta).$$

Thus $w(a)\cdot a \geq w(b)\cdot a$. By (19), $w(a)=w(b)$. The equality (20) then holds for any $\varepsilon \geq 0$.

Case 2. $x\cdot v(a) > 0 > x\cdot v(b)$ and $x\cdot(w(b)-w(a)) > 0$ for some $x\in\mathbb{R}^S$. Take a sufficiently small $\delta > 0$ and $c=\delta x$ such that

$$\begin{aligned} V(a+c) &= V(a)+\delta(x\cdot v(a))+o(\delta) > V(a)\\ V(b+c) &= V(b)+\delta(x\cdot v(b))+o(\delta) < V(b)\\ W(a+c)-W(b+c) &= \delta(x\cdot w(a))-\delta(x\cdot w(b))+o(\delta) < 0. \end{aligned}$$

This contradicts (21). □

If $W=V$, then $\succeq$ is extremely cautious, $\Pi=\Delta$, and the utility representation (4) holds for $\varepsilon = 1$ and any $p\in\Delta$.

Lemma 3. *If $W\neq V$, then there are unique $0\leq\varepsilon<1$ and $p\in\Delta$ such that*

$$W(a) = \varepsilon V(a)+(1-\varepsilon)p\cdot a$$

for all $a\in\mathbb{R}^S$. Moreover, p is the only probability measure in $\mathscr{P}$ such that for all $f,g\in\mathscr{H}$, if f is more secure than g and $f(p)\succ g(p)$, then $f\succ g$.

Proof. Suppose that $W\neq V$. If $v(a)=p$ is constant for all $a\in\mathbb{D}$, then $V(a)=p\cdot a$ for all $a\in\mathbb{D}$, and by continuity, for all $a\in\mathbb{R}^S$. Then the inequality

$$\min_{q\in\Pi} q\cdot a = W(a) \geq p\cdot a \qquad \text{for all } a\in\mathbb{R}^S$$

implies that $\Pi=\{p\}$, which contradicts $W\neq V$.

Thus, v is not constant on $\mathbb{D}$, and there are $b,c\in\mathbb{D}$ such that $v(b)\neq v(c)$. By Lemma 2, there is $\varepsilon\geq 0$ such that

$$w(b)-w(c) = \varepsilon(v(b)-v(c)). \tag{22}$$

Take any $a\in\mathbb{D}$. I claim that

$$w(a) = \varepsilon v(a)+\hat{p}, \tag{23}$$

where $\hat{p} = w(b)-\varepsilon v(b) = w(c)-\varepsilon v(c)$. To show this claim, let

$$\begin{aligned} B &= \{w(b)+\gamma(v(a)-v(b)) : \gamma\geq 0\}\\ C &= \{w(c)+\gamma(v(a)-v(c)) : \gamma\geq 0\}. \end{aligned}$$

If $v(a)=v(b)$ or $v(a)=v(c)$, then B or, respectively, C is a singleton. If $v(a)\neq v(b)$ and $v(a)\neq v(c)$, then B and C are rays in $\mathbb{R}^S$. Moreover, the directions of these rays, $v(a)-v(b)$ and $v(a)-v(c)$ respectively, are linearly independent because $v(a)$, $v(b)$, and $v(c)$ are distinct extreme points in Δ. Therefore, the rays B and C have at most one point in common. However, $\varepsilon v(a)+\hat{p} \in B\cap C$ for $\gamma=\varepsilon$, and by Lemma 2, $w(a)\in B\cap C$. It follows that $w(a)=\varepsilon v(a)+\hat{p}$.

By (18), (19), and (23),

$$W(a)=w(a)\cdot a=\varepsilon v(a)\cdot a+\hat{p}\cdot a=\varepsilon V(a)+\hat{p}\cdot a$$

for all $a\in\mathbb{D}$. Rockafellar (1970, Theorem 25.5) shows that the complement of $\mathbb{D}$ has measure zero, and hence $\mathbb{D}$ is *dense* in $\mathbb{R}^S$. By continuity, for all $a\in\mathbb{R}^S$,

$$W(a)=\varepsilon V(a)+\hat{p}\cdot a.$$

By (17),

$$\hat{p}\cdot a\geq(1-\varepsilon)V(a). \tag{24}$$

I show that $\varepsilon<1$ and $\hat{p}=(1-\varepsilon)p$ for some $p\in\Delta$. Consider three cases.

(i) $\varepsilon>1$. Recall that there exist two distinct points $v(b), v(c)\in\Delta$. Let $a=v(b)-v(c)$. Then $V(a)+V(-a)<0$ because

$$\begin{aligned}V(a)&\leq v(c)\cdot a<v(b)\cdot a\\ V(-a)&\leq -v(b)\cdot a<v(c)\cdot a.\end{aligned}$$

On the other hand, by (24), $V(a)+V(-a)\geq\frac{\hat{p}}{1-\varepsilon}\cdot a+\frac{\hat{p}}{1-\varepsilon}\cdot(-a)=0$, which is a contradiction.

(ii) $\varepsilon=1$. By (24), $W=V$, which contradicts $W\neq V$.

(iii) $\varepsilon<1$. Let $p=\hat{p}/(1-\varepsilon)$. By (24), $p\cdot a\geq V(a)=\min_{q\in\Delta} q\cdot a$ for all $a\in\mathbb{R}^S$. As Δ is convex and closed, $p\in\Delta$ by the separating hyperplane argument.

I now turn to the uniqueness part. The parameter $0\leq\varepsilon<1$ is uniquely determined by (22) and $p=\hat{p}/(1-\varepsilon)$ is unique.

Moreover, for any $p'\in\mathscr{P}$ such that $p'\neq p$, there are acts $f,g\in\mathscr{H}$ such that

$$p'\cdot u(f)>p'\cdot u(g),\quad p\cdot u(f)\leq p\cdot u(g),\text{ and } V(u(f))=V(u(g)). \tag{25}$$

To construct such f and g, take an event $A\subset S$ such that $p'(A)>p(A)$. Let $\pi_*(A)=\min_{q\in\Delta} q(A)$ and $\pi_*(A^c)=\min_{q\in\Delta} q(A^c)$. Take vectors $a,b\in\mathbb{R}^S$ such that

$$a_s=\begin{cases}1-\pi_*(A) & \text{if } s\in A\\ -\pi_*(A) & \text{if } s\in A^c\end{cases}\quad\text{and}\quad b_s=\begin{cases}-\pi_*(A^c) & \text{if } s\in A\\ 1-\pi_*(A^c) & \text{if } s\in A^c.\end{cases}$$

By construction, $p'\cdot a>p\cdot a$, $p\cdot b>p'\cdot b$, $p\cdot a\geq V(a)=0$, and $p\cdot b\geq V(b)=0$. If $p\cdot a=p\cdot b$, then take $f,g\in\mathscr{H}$ such that $u(f)=a$ and $u(g)=b$. If $p\cdot a\neq p\cdot b$, then take $f,g\in\mathscr{H}$ such that $u(f)=(p\cdot b)a$ and $u(g)=(p\cdot a)b$.

It follows from (25) that for any $p'\in\mathscr{P}$ such that $p'\neq p$ there are $f,g\in\mathscr{H}$ such that $f(p')\succ g(p')$, f is more secure than g, but still $g\succeq f$. The proof of Lemma 3 is complete. □

Lemma 3 delivers the required utility representation (4) for the preference $\succeq$. Moreover, it implies that if $\succeq$ is not extremely cautious, then this representation is unique up to a positive linear transformation of u, and p is the only probability measure that satisfies condition (9), in Theorem 2.

Now turn to Theorem 3. Let $\succeq'$ be another preference that has a utility representation (4) with a tuple $(u', \Delta', p', \varepsilon')$. Suppose that $\succeq'$ is more averse to interim ambiguity than $\succeq$. Then for all lotteries $l, l' \in \mathscr{L}$, $l \succeq l'$ implies that l is more secure than l', and hence $l \succeq' l'$. Thus, u' and u are equal up to a positive linear transformation. Without loss of generality, let $u' = u$. If $\varepsilon' = 1$, then $\varepsilon' \geq \varepsilon$ and p' can be taken equal to p. Assume that $\varepsilon' < 1$ and $p' \neq p$. Take acts f, g that satisfy (25). Then g is more secure than f, $g \succeq f$, but $f \succ' g$, which violates (10). Thus $p' = p$. If $\Delta = \{p\}$ is a singleton, then ε' can be taken equal to ε. Assume that Δ is not a singleton, and $\varepsilon > \varepsilon'$. Then the function v is not constant on $\mathbb{D}$, and hence there is a such that $v(a) \neq p$. As $\varepsilon > \varepsilon'$, the point $w'(a)$ does not belong to the segment $[v(a), w(a)]$. By the separating hyperplane theorem, there is $x \in \mathbb{R}^S$ such that $v(a) \cdot x > 0$, $w(a) \cdot x > 0$, but $w'(a) \cdot x < 0$. Then for sufficiently small $\delta > 0$, $V(a + \delta x) > V(a)$, $W(a + \delta x) > W(a)$, but $W'(a + \delta x) < W'(a)$. Take acts f, g such that $u(f) = a + \delta x$ and $u(g) = a$. (If needed, rescale a and x so that $a, a + \delta x \in [-1, 1]^S$.) Then f is more secure than g, $f \succ g$, but $g \succ' f$, which violates (10). Thus $\varepsilon' \geq \varepsilon$.

To extend the utility representation (4) to an arbitrary state space (S, Σ), consider two cases.

Case 1. For all events $A \in \Sigma$, $\pi(A) = \pi_*(A)$. Then for any finite subalgebra $\Sigma' \subset \Sigma$, Theorem 1 implies that the utility function $U(f) = \min_{q \in \Delta} f(q)$ represents the preference $\succeq$ restricted to Σ' measurable acts. Thus, U represents $\succeq$ on all of $\mathscr{H}$.

Case 2. There exists an event $A \in \Sigma$ such that $\pi(A) > \pi_*(A)$. Let

$$\varepsilon = \frac{1 - \pi(A) - \pi(A^c)}{1 - \pi_*(A) - \pi_*(A^c)}.$$

For all events $B \in \Sigma$, let $p(B) = (\pi(B) - \varepsilon \pi_*(B))/(1 - \varepsilon)$. Then for any finite subalgebra $\Sigma' \subset \Sigma$ such that $A \in \Sigma'$, the finite case of Theorem 1 implies that p is finitely additive on Σ' and the preference $\succeq$ is represented by (4) with components $(u, \varepsilon, \Delta, p)$ when restricted to Σ' measurable acts. Thus, p is finitely additive on all of Σ, and the preference $\succeq$ has the utility representation (4) with components $(u, \varepsilon, \Delta, p)$ on all of $\mathscr{H}$.

B. Robustness of p and ε

Suppose that for all events $A \in \Sigma$, the observed values $\hat{\pi}(A)$ and $\hat{\pi}_*(A)$ of the subjective willingness to bet on any event A are generated by

$$\begin{aligned} \hat{\pi}_*(A) &= \min_{q \in \Delta} q(A) + e_*(A) \\ \hat{\pi}(A) &= \varepsilon \min_{q \in \Delta} q(A) + (1 - \varepsilon) p(A) + e(A) \end{aligned} \tag{26}$$

where $\varepsilon \in [0, 1]$, $\Delta \in \mathscr{C}$, $p \in \Delta$ are the sought-after true components of the model and $e_*(A)$ and $e(A)$ are error terms that do not exceed in magnitude a given threshold $e > 0$.

It seems plausible that e should be small when the decision maker has a monetary incentive not to postpone her ex ante decisions, but this incentive is minor compared to the monetary stakes x and y.

For each event A, let $\hat{E}(A)$ be the range of parameters $\hat{\varepsilon}$ that are consistent with the observed values $\hat{\pi}(A)$, $\hat{\pi}_*(A)$, $\hat{\pi}(A^c)$, and $\hat{\pi}_*(A^c)$ and specification (26) so that

$$\begin{aligned} |q_*(A)-\hat{\pi}_*(A)| &\le e \\ |(1-q^*(A))-\hat{\pi}_*(A^c)| &\le e \\ |\hat{\varepsilon}q_*(A)+(1-\hat{\varepsilon})\hat{p}(A)-\hat{\pi}(A)| &\le e \\ |\hat{\varepsilon}(1-q^*(A))+(1-\hat{\varepsilon})\hat{p}(A)-\hat{\pi}(A^c)| &\le e \end{aligned} \tag{27}$$

for some $0\le q_*(A)\le \hat{p}(A)\le q^*(A)\le 1$. In particular, these inequalities hold for the true parameter value ε and $q_*(A)=\min_{q\in\Delta} q(A)\le \hat{p}(A)=p(A)\le q^*(A)=\max_{q\in\Delta} q(A)$. Therefore, ε must belong to $\hat{E}(A)$. Let

$$\hat{E}=\cap_{A\in\Sigma}\hat{E}(A).$$

If S or Σ is finite, which must be the case in any empirical study, then the construction of $\hat{E}$ is a (tedious) arithmetic exercise that can be done by a computer.

Note that for any A such that $1-\hat{\pi}_*(A)-\hat{\pi}_*(A^c)>2e$, the length of $\hat{E}(A)$ does not exceed

$$\begin{aligned} &\frac{1-\hat{\pi}(A)-\hat{\pi}(A^c)+2e}{1-\hat{\pi}_*(A)-\hat{\pi}_*(A^c)-2e}-\frac{1-\hat{\pi}(A)-\hat{\pi}(A^c)-2e}{1-\hat{\pi}_*(A)-\hat{\pi}_*(A^c)+2e}= \\ &\qquad\qquad \frac{4e(1-\hat{\pi}(A)-\hat{\pi}(A^c))}{(1-\hat{\pi}_*(A)-\hat{\pi}_*(A^c))^2-4e^2}\le\frac{4e}{1-\hat{\pi}_*(A)-\hat{\pi}_*(A^c)-2e}. \end{aligned}$$

Therefore, for any $\hat{\varepsilon}\in\hat{E}$,

$$|\hat{\varepsilon}-\varepsilon|\le\min_{A\in\Sigma}\frac{4e}{1-\hat{\pi}_*(A)-\hat{\pi}_*(A^c)-2e}. \tag{28}$$

In particular, if both $\hat{\pi}_*(A)$ and $\hat{\pi}_*(A^c)$ are close to zero for some event A, then any value $\hat{\varepsilon}\in\hat{E}$ estimates ε with an error not exceeding $4e/(1-2e)$.

Finally, for all $A\in\Sigma$ and $\hat{\varepsilon}\in\hat{E}$, it follows from (26) and (27) that

$$\begin{gathered} \Big|q_*(A)-\min_{q\in\Delta} q(A)\Big|\le 2e \\ \Big|\hat{\varepsilon}q_*(A)+(1-\hat{\varepsilon})\hat{p}(A)-\varepsilon\min_{q\in\Delta} q(A)-(1-\varepsilon)p(A)\Big|\le 2e \end{gathered}$$

and hence $\Big|(\varepsilon-\hat{\varepsilon})(p(A)-\min_{q\in\Delta} q(A))+(1-\hat{\varepsilon})(\hat{p}(A)-p(A))\Big|\le 4e$. Without loss in generality, $\varepsilon-\hat{\varepsilon}$ and $\hat{p}(A)-p(A)$ have the same sign. (Otherwise, replace A with A^c, $p(A)$ with $1-p(A)$, and $\hat{p}(A)$ with $1-\hat{p}(A)$.) It follows that $\Big|(1-\hat{\varepsilon})(\hat{p}(A)-p(A))\Big|\le 4e$ and hence

$$\Big|\hat{p}(A)-p(A)\Big|\le\frac{4e}{1-\hat{\varepsilon}}. \tag{29}$$

The inequalities (28) and (29) suggest that $\hat{\varepsilon}$ and $\hat{p}(A)$ are robust estimates for the parameter ε and probabilities $p(A)$ in specification (26), except for the boundary cases when ε is close to one or Δ is small (then $1-\hat{\pi}_*(A)-\hat{\pi}_*(A^c)$ is small for all $A\in\Sigma$). Note that the estimated function $\hat{p}$ need not be additive. If additivity is added to the constraints (27), then the quality of the estimate can improve further, but this improvement is hard to evaluate.

References

Anscombe, Francis J. and Robert J. Aumann (1963), "A definition of subjective probability." *Annals of Mathematical Statistics*, 34, 199–205. [200, 202]

Becker, Selwyn W. and Fred O. Brownson (1964), "What price ambiguity? or the role of ambiguity in decision-making." *Journal of Political Economy*, 72, 62–73. [202]

Berger, James O. (1994), "An overview of robust Bayesian analysis." *Test*, 3, 5–124. [201]

Bewley, Truman F. (2002), "Knightian decision theory. Part I." *Decisions in Economics nad Finance*, 25, 79–110. [200]

Camerer, Colin and Martin Weber (1992), "Recent developments in modeling preferences: Uncertainty and ambiguity." *Journal of Risk and Uncertainty*, 5, 325–370. [202]

Carlier, Guillaume, Rose-Anne Dana, and Niousha Shahidi (2003), "Efficient insurance contracts under epsilon-contaminated utilities." *Geneva Papers on Risk and Insurance Theory*, 28, 59–71. [202, 213]

Chateauneuf, Alain, Jurgen Eichberger, and Simon Grant (2007), "Choice under uncertainty with the best and worst in mind: Neo-additive capacities." *Journal of Economic Theory*, 137, 538–567. [215]

Danan, Eric and Anthony Ziegelmeyer (2008), "Are preferences complete? An experimental measurement of indecisiveness under risk." Unpublished paper, Thema, Université de Cergy-Pontoise. [204]

Dekel, Eddie, Barton L. Lipman, and Aldo Rustichini (2001), "Representing preferences with a unique subjective state space." *Econometrica*, 69, 891–934. [199]

Dhar, Ravi (1997), "Consumer preference for a no-choice option." *Journal of Consumer Research*, 24, 215–231. [199]

Eichberger, Jurgen and David Kelsey (1999), "E-capacities and the Ellsberg paradox." *Theory and Decision*, 46, 107–140. [202]

Ellsberg, Daniel (1961), "Risk, ambiguity, and the Savage axioms." *Quarterly Journal of Economics*, 75, 643–669. [199, 202]

Epstein, Larry G., Massimo Marinacci, and Kyoungwon Seo (2008), "Coarse contingencies and ambiguity." *Theoretical Economics*, 2, 355–394. [214]

Epstein, Larry G. and Tan Wang (1994), "Intertemporal asset pricing under Knightian uncertainty." *Econometrica*, 62, 283–322. [202, 213]

Fox, Craig R. and Amos Tversky (1995), "Ambiguity aversion and comparative ignorance." *Quarterly Journal of Economics*, 110, 585–603. [205]

Gajdos, Thibault, Takashi Hayashi, Jean-Marc Tallon, and Jean-Christophe Vergnaud (2008), "Attitude toward imprecise information." *Journal of Economic Theory*, 140, 27–65. [202]

Ghirardato, Paolo, Fabio Maccheroni, and Massimo Marinacci (2004), "Differentiating ambiguity and ambiguity attitude." *Journal of Economic Theory*, 118, 133–173. [201, 206, 209, 217]

Ghirardato, Paolo, Fabio Maccheroni, and Massimo Marinacci (2005), "Certainty independence and the separation of utility and beliefs." *Journal of Economic Theory*, 120, 129–136. [206]

Ghirardato, Paolo and Massimo Marinacci (2002), "Ambiguity made precise: A comparative foundation." *Journal of Economic Theory*, 102, 251–289. [211]

Gilboa, Itzhak (1988), "A combination of expected utility and maximin decision criteria." *Journal of Mathematical Psychology*, 32, 405–420. [202]

Gilboa, Itzhak, Fabio Maccheroni, Massimo Marinacci, and David Schmeidler (2008), "Objective and subjective rationality in a multiple prior model." Unpublished paper. [202, 217]

Gilboa, Itzhak and David Schmeidler (1989), "Maxmin expected utility with non-unique prior." *Journal of Mathematical Economics*, 18, 141–153. [201]

Herstein, Israel N. and John Milnor (1953), "An axiomatic approach to measurable utility." *Econometrica*, 21, 291–297. [209]

Hodges, Joseph L., Jr. and Erich L. Lehmann (1952), "The use of previous experience in reaching statistical decisions." *Annals of Mathematical Statistics*, 23, 396–407. [201]

Jaffray, Jean-Yves (1988), "Choice under risk and the security factor: An axiomatic model." *Theory and Decision. An International Journal for Philosophy and Methodology of the Social Sciences*, 24, 169–200. [202]

Koopmans, Tjalling C. (1964), "On flexibiliy of future preference." In *Human Judgments and Rationality* (II Shelley, Maynard W. and Glenn L. Bryan, eds.), 243–254, Wiley. [199]

Kopylov, Igor (2008), "Subjective probability and confidence." Unpublished paper, Department of Economics, University of California, Irvine. [215]

Kreps, David M. (1979), "A representation theorem for 'preference for flexibility'." *Econometrica*, 47, 565–577. [199]

Kreps, David M. (1988), *Notes on the Theory of Choice*. Westview Press, Boulder. [206]

Maccheroni, Fabio, Massimo Marinacci, and Aldo Rustichini (2006), "Ambiguity aversion, robustness, and the variational representation of preferences." *Econometrica*, 74, 1447–1498. [206]

Machina, Mark J. and David Schmeidler (1992), "A more robust definition of subjective probability." *Econometrica*, 60, 745–780. [211]

Nishimura, Kiyohiko G. and Hiroyuki Ozaki (2004), "Search and Knightian uncertainty." *Journal of Economic Theory*, 119, 299–333. [202, 213]

Nishimura, Kiyohiko G. and Hiroyuki Ozaki (2006), "An axiomatic approach to epsilon-contamination." *Economic Theory*, 27, 333–340. [202]

Rockafellar, R. Tyrrell (1970), *Convex Analysis*. Princeton University Press, Princeton, New Jersey. [211, 220]

Sen, Amartya (1988), "Freedom of choice: Concept and content." *European Economic Review*, 32, 269–294. [204]

Tversky, Amos and Eldar Shafir (1992), "Choice under conflict: The dynamics of deferred decision." *Psychological Science*, 3, 358–361. [199]

Submitted 2008-10-20. Final version accepted 2009-3-13. Available online 2009-3-13.