SUBJECTIVE PROBABILITY, CONFIDENCE, AND BAYESIAN UPDATING

Igor Kopylov*

Institute for Mathematical Behavioral Sciences, and

Department of Economics, University of California, Irvine, CA 92697

Abstract

I define subjective probabilities for an *ambiguity averse* agent who is given an exogenous *information set* Δ that must contain the true probability law on the state space S. The agent in my model evaluates every uncertain prospect f via a mixture of the least favorable scenario in Δ and her *unique subjective belief* p. The weights in this mixture are unique and reveal the agent's *confidence* in her belief p. To characterize this model, I take four standard axioms—order, continuity, monotonicity, and Independence—and use the information set Δ to modify the last two conditions in this list. Moreover, I relax the well-known *dynamic consistency* principle and characterize the *Bayesian updating* rule for the belief p conditional on any non-null event $E \subset S$.

JEL Classification Numbers: D81

Keywords: Ambiguity; Bayesian Updating; Dynamic Consistency; Subjective Probability; Confidence; Multiple Priors Model; epsilon contamination

1 Introduction

In the classic paradigm of Ramsey [33], de Finetti [9], Savage [35], and Machina and Schmeidler [29], *subjective probabilities* over all events in the state space S are derived uniquely from observable preferences over monetary bets and other uncertain prospects rather than from objective symmetries or frequencies. *Bayesian updating* for subjective probabilities conditional on any non-null event in S is then implied by the *dynamic consistency* principle (see Ghirardato [17]).

Yet these models do not accommodate *ambiguity averse* agents who can persistently bet on some event E rather than A, but also on the complement $E^c = S \setminus E$

*I thank Mark Machina, Larry Epstein, Peter Wakker, Marciano Siniscalchi, RUD participants, and anonymous referees for their comments.

rather than $A^c = S \setminus A$. As first noted by Ellsberg [10], this betting preference is common when the events E and E^c are known to be equally likely, but A and A^c have unknown or vague objective probabilities. In the three-color version of the Ellsberg Paradox, the state of the world $s \in S = \{R, G, B\}$ is determined by the color—reg, green or blue—of a ball drawn randomly from an urn, the proportion of red balls in the urn is known to be $\frac{1}{3}$, but the proportions of green and blue balls are unknown. In this case, it is typical to bet on $E = \{R\}$ and $E^c = \{G, B\}$ rather than $A = \{B\}$ and $A^c = \{R, G\}$ respectively.

More broadly, ambiguity aversion is plausible when agents are informed that the true probability law p_0 on the state space S must belong to a given set Δ, but no other constraints on p_0 are specified. The exogenous *information set* Δ is natural in many experimental and theoretical settings. Starting from Becker and Brownson [2], there have been at least forty empirical studies listed by Oechssler and Roomets [32, p.3]. in the Ellsberg-style frameworks where objects (e.g. balls, cards) were drawn at random from compositions (e.g. urns, decks) that were not fully disclosed to the subjects. Such partial information can be easily translated into information sets (see various examples in Chew, Miao, and Zhong [7]). In theoretical literature, information sets have been used in models of unambiguous events (Epstein and Zhang [14, p.269]), smooth ambiguity (Klibanoff, Marinacci, and Mukerji [25, p.1860]), objective rationality (Gilboa, Maccheroni, Marinacci, and Schmeidler [20]), robust Bayesian analysis (reviewed by Berger [3]), and other applications.

In this paper, I define a *unique subjective probabilistic belief* p in terms of ambiguity averse preferences and an exogenous information set Δ. Moreover, I formulate a weak form of dynamic consistency that implies *Bayesian updating* for p and can be combined with various updating rules for the information set Δ.

Formally, I model preferences $\succeq$ over Anscombe–Aumann's [1] *acts* f that map states of the world $s \in S$ into lotteries $f(s)$—objective distributions over deterministic prizes. Given any act f and probability measure q on S, let $f(q)$ denote the lottery *induced* by f via q. In particular, if S is finite, let

$$f(q) = \sum_{s \in S} q(s) f(s).$$

My first result (Theorem 1) axiomatizes the utility representation

$$U(f) = (1 - \varepsilon) u\left(f(p)\right) + \varepsilon \min_{q \in \mathrm{cl}(\mathrm{co}\, \Delta)} u\left(f(q)\right), \tag{1.1}$$

where $\varepsilon \in [0, 1]$, p is a probability measure within the closed convex hull of the information set Δ, and u is a vNM expected utility index over lotteries. This representation is unique up to a positive linear transformation of u (except that p is arbitrary when $\varepsilon = 1$.)

Moreover, the decision maker's *subjective belief* p in (1.1) is shown to be the only probability measure that satisfies

$$f(p) \succ g(p) \quad \Rightarrow \quad f \succ g \tag{1.2}$$

for all acts f and g such that f is *more secure* than g with respect to the least favorable scenarios in the information set. By contrast, condition (1.2) need not hold when g is more secure than f. In this case, ambiguity aversion can motivate the preference $g \succeq f$ even if the agent holds the opposite ranking $f(p) \succ g(p)$ for the lotteries induced by f and g via her subjective belief p. Such preference reversals contradict Machina and Schmeidler's [29] definition of *probabilistic sophistication*, but still agree with the more general use of subjective probabilities in my model.

To characterize (1.1), I take four standard axioms (e.g. as in Fishburn [15, p. 178])—Order, Continuity, Monotonicity, and Independence—and use the extra primitive Δ to modify the last two conditions in this list. To accommodate ambiguity aversion, preference reversals $f \succeq g$ and $\alpha g + (1-\alpha)h \succ \alpha f + (1-\alpha)h$ are allowed when the mixture $\alpha g + (1-\alpha)h$ *hedges* ambiguity in Δ better than $\alpha h + (1-\alpha)h$ does.[1]

My other results (Theorems 2 and 3 below) use violations of Independence to characterize the parameter ε, or equivalently $1-\varepsilon$, in terms choice behavior. These results suggest roughly that the decision maker's *confidence* in her belief p is revealed by her willingness to comply with the Independence axiom.

An important special case of representation (1.1) is obtained when the decision maker has no information about objective probabilities, and hence, Δ equals the set $\mathcal{P}$ of all probability measures on the state space S. In this case of *complete ignorance*, representation (1.1) takes the form

$$U(f) = (1-\varepsilon)u\left(f(p)\right) + \varepsilon \min_{s \in S} u\left(f(s)\right). \tag{1.3}$$

In this case, the subjective probability measure p represents the betting preferences over all events $A \neq S$. Thus, if the state space S is sufficiently rich, the probability measure p can be elicited from the comparative likelihood relation as in de Finetti [9] or Kopylov [27].

1.1 Bayesian Updating and Dynamic Consistency

To model updating of subjective probabilities on any given event E, assume that two preferences $\succeq$ and $\succeq_E$ are observed before and after E is learnt by the decision maker.

[1]The formal definitions of comparative security levels and hedging are discussed later.

Given any probability measure q such that $q(E) > 0$, its Bayesian update $q|E$ conditional on E is defined by the formula

$$(q|E)(A) = \frac{q(A \cap E)}{q(E)}$$

for all events $A \subset S$. This formula can be also applied to the exogenous information set Δ via

$$\Delta|E = \{q|E : q \in \Delta \text{ and } q(E) > 0\}.$$

Then Theorem 1 can be used to characterize the utility representations

$$\begin{aligned} U(f) &= (1-\varepsilon)u\left(f(p)\right) + \varepsilon \min_{q \in \mathrm{cl}(\mathrm{co}\,\Delta)} u\left(f(q)\right), \\ U_E(f) &= (1-\lambda)u\left(f(p_E)\right) + \lambda \min_{q \in \mathrm{cl}(\mathrm{co}\,\Delta|E)} u\left(f(q)\right) \end{aligned}$$

where $\varepsilon, \lambda \in [0,1]$, p and p_E are probability measures within the closed convex hulls of Δ and $\Delta|E$ respectively.

My last result (Theorem 4) characterizes the Bayesian updating for the subjective beliefs $p_E = p|E$ via a weak form of *dynamic consistency.* The implication

$$f \succeq g \quad \Rightarrow \quad f \succeq_E g \tag{1.4}$$

is required for all acts f and g such that $f = g$ on E^c, f is less secure than g, but becomes more secure than g after E is learnt.

This result shows that Bayesian updating of subjective beliefs can be derived from weaker behavioral implications for ambiguity averse agents than for subjective expected utility maximizers, for whom implication (1.4) must hold for all acts f and g (e.g. Ghirardato [17]). For example, recall the three-color Ellsberg urn setting with $S = \{R, G, B\}$ and $\Delta = \{q \in \mathcal{P} : q(R) = \frac{1}{3}\}$. Let $E = \{R, G\}$, and consider the following uncertain prospects

	f	g	f'	g'
R	0	100	$100 + M$	0
G	100	0	0	100
B	100	100	100	100

with payoffs given in risk neutral utils. Then the typical ambiguity averse preference $f \succ g$ is motivated by the fact that f is more secure than g and hence, need not satisfy (1.4). By contrast, the weak dynamic consistency (1.4) will apply if preference $f' \succ g'$ holds for some $M \in [0, 100]$. In this case f' is less secure than g because $\frac{1}{3}(100 + M) \leq \frac{2}{3}100$, but becomes more secure than g after E is learnt because $\frac{1}{3}(100 + M) > 0$.

1.2 Related Literature

1.2.1 Epsilon Contamination in the Multiple Priors Model

Obviously, representation (1.1) is a special case of Gilboa–Schmeidler's [21] maxmin expected utility (MEU)

$$U(f) = \min_{q \in \Pi} u\left(f(q)\right), \tag{1.5}$$

where the set of priors

$$\Pi = (1-\varepsilon)\{p\} + \varepsilon \mathrm{cl}(\mathrm{co}\,\Delta)$$

has the added parametric structure of *epsilon contamination.* Conversely, any representation (1.1) can be rewritten in the form (1.5) for various pairs (ε, Δ), most simply for $\varepsilon = 1$ and $\Delta = \Pi$.

Ellsberg [10, p.663–669] suggests the functional form (1.1) 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'. In an early experimental study, Becker and Brownson [2] find some evidence that people put a constant weight ε on the least favorable probabilistic scenario in the information set. (This study estimates the average weight ε to be 0.768.) Other experiments (reviewed by Camerer and Weber [5]) produce mixed results.

The epsilon contamination structure (1.1) has many applications in statistics, decision theory, and economics. Following Hodges and Lehmann [24], this functional form has been used in robust Bayesian analysis (e.g. Berger and Berliner [4], Moreno and Cano [30]). The complete ignorance model (1.3) has been applied to model asset pricing (Epstein and Wang [13]), search (Nishimura and Ozaki [31]), and insurance (Carlier, Dana, Shahdi [6].)

Gajdos, Hayashi, Tallon, and Vergnaud [16] derive epsilon 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). All of the above models impose lists of axioms much different from those adopted here.

Kopylov [26] formulates a representation result which is similar to Theorem 1, but has distinct primitives. Instead of taking Δ exogenous, this set is derived endogenously from an incomplete binary relation $\succeq_*$ that represents choices that the decision maker is willing to make immediately without waiting for the objective probability law p_0 to become clear. The endogenous definition of Δ has its own practical limitations, as it requires an additional decision stage after the true probability law has been announced. The axioms and interpretation for representation (1.1) in the choice deferral model are distinct from the exogenous case as well. More substantially, my characterizations of Bayesian updating of subjective beliefs (Theorem 4 and Proposition 5) and the comparative confidence result (Theorem 3) have no counterparts or analogies in Kopylov [26].

1.2.2 Bayesian Updating Under Ambiguity Aversion

The Bayesian updating of the set Π in the multiple priors model via the prior-by-prior rule

$$\Pi|E = \{q|E : q \in \Pi \text{ and } q(E) > 0\} \tag{1.6}$$

is characterized by Ghirardato, Maccheroni, and Marinacci (GMM) [18]. They impose dynamic consistency (1.4) only when $f \succeq g$ is unambiguous, that is, $f(q) \succeq g(q)$ for all $q \in \Pi$.

The prior-by-prior model (1.6) differs from mine in several respects. First, it does not require the information set Δ as a primitive. Second, if Δ is given and $\Pi == \varepsilon\Delta + (1-\varepsilon)\{p\}$, then $\Pi|E$ can be uniquely rewritten in the form $\lambda(\Delta|E) + (1-\lambda)\{p_E\}$, but p_E need not equal $p|E$.

For example, let $S = \{R, G, B\}$ and $\Delta = \{q \in \mathcal{P} : q(R) = \frac{1}{3}\}$. Let $E = \{R, G\}$, $p(R) = p(G) = p(B) = \frac{1}{3}$, and $\varepsilon = \frac{1}{2}$. Then

$$\Pi|E = (\varepsilon\Delta + (1-\varepsilon)\{p\})\,|E = \tfrac{2}{5}(\Delta|E) + \tfrac{3}{5}\{p_E\}$$

where $p_E(R) = \frac{4}{9} \neq \frac{1}{2} = (p|E)(R)$.

In a more practical vein, my model allows the agent's confidence parameter $1-\lambda$ to vary upon the observed event E. For example, it accommodates agents who

(i) conform to expected utility maximization with $\lambda = 0$ once they observe an event E that provides sufficient support for their subjective belief p,

(ii) become less confident (i.e. more ambiguity averse) in response to surprising events E that have relatively low subjective probabilities $p(E)$.

The prior-by-prior updating in GMM is more rigid and makes initial ambiguity aversion persistent regardless of how well the subjective belief p agrees with the observed events E.

Sinischalchi [36] characterizes the vector expected utility (VEU) representation

$$U(f) = E_p[u \circ f] + A(E_p[\xi_i \cdot u \circ f])$$

where (i) the baseline prior p is subjective and updated through the Bayesian rule, (ii) the adjustment factors ξ_i for $i = 1 \ldots n$ are random variables such that $E_p(\xi_i) = 0$, and (iii) the functional A is such that $A(-\phi) = A(\phi)$ for all vectors

$\phi \in \mathbb{R}^n$. This representation overlaps with the multiple priors model only when p is the center of symmetry for the set Π. Note that even if the information set Δ is symmetric, p need not be its center in my model. Similarly to the general MEU, Sinischalchi's VEU representation is not parametric and hence, should be more difficult to calibrate than the epsilon contamination structure.

Other ways to incorporate various forms of dynamic consistency into models of ambiguity aversion include the *rectangularity* condition on the set Π (Epstein and Schneider [12]), updating priors in a menu-dependent way (Hanany and Klibanoff [23]), and changing the subjective perception of decision problems (Halevy and Feltkamp [22]).

2 Preliminaries

Adopt a version of Anscombe–Aumann's [1] decision framework. Given are a set X of *outcomes*, a set S of *states* of nature, and an algebra $\Sigma \subset 2^S$ of *events*. Based on these primitives, define

- the set $\mathcal{L} = \{l, \dots\}$ of all *lotteries*—probability measures on X with finite support,
- the set $\mathcal{U}$ of all expected utility functions on $\mathcal{L}$,
- the set $\mathcal{P} = \{p, q, \dots\}$ of all finitely additive probability measures on (S, Σ) with the weak* topology,[2]
- the set $\mathcal{H} = \{f, g, \dots\}$ of all *acts*—Σ-measurable functions $f : S \to \mathcal{L}$ that have a finite range in $\mathcal{L}$.

Endow the set $\mathcal{H}$ with a natural *mixture* operation: for any $f, g \in \mathcal{H}$ 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 $l \in \mathcal{L}$. Given any acts $f, g \in \mathcal{L}$ and any event $A \in \Sigma$, define a composite act

$$fAg = \begin{cases} f(s) & \text{if } s \in A \\ g(s) & \text{if } s \notin A. \end{cases}$$

Interpret any act $f \in \mathcal{H}$ 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.

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

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

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

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

Suppose that there is an objective probability distribution $p_0 \in \mathcal{P}$ on the state space, but the decision maker does not know p_0 precisely. To describe this ambiguity, take a non-empty *information set* $\Delta \subset \mathcal{P}$ as another primitive of the model. Assume that the decision maker knows only that p_0 belongs to the set Δ.

For example, Δ can be specified as follows.

(i) Let $\Delta = \mathcal{P}$. This case captures *complete ignorance* as the decision maker is assumed to know nothing about p_0.

(ii) Let

$$\Delta = \{q \in \mathcal{P} : q(A) = p_0(A) \quad \text{for all } A \in \Gamma\}$$

if objective probabilities $p_0(A)$ are known only for events A in a subclass $\Gamma \subset \Sigma$. Epstein and Zhang [14, p. 269] use this specification to motivate their model of *unambiguous* events. This structure also accommodates applications (e.g. Moreno and Cano [30]) where only the median or some other percentiles of the true probability law are known.

(iii) Let

$$\Delta = \{q \in \mathcal{P} : q(A) \geq \nu(A) \quad \text{for all } A \in \Sigma\}$$

if the decision maker knows only some lower bounds $\nu(A)$ for objective probabilities $p_0(A)$. (Note that imposing an upper bound on $p_0(A)$ is equivalent to imposing a lower bound on $p_0(A^c)$.) This specification accommodates various empirical settings (such as Becker and Brownson [2], Curley and Yates [8]) that have been used to test individual response to ambiguity. In these settings, subjects were confronted with bets on events for which only intervals of possible probabilities were given.

(iv) Let Δ be a parametric class of probability measures (e.g. uniform, normal, Poisson, binomial etc.) on a suitable state space S. This structure is common in robust Bayesian analysis (e.g. Berger and Berliner [4]). In this case, the set Δ is often not even convex. While the lack of convexity (and closedness) is convenient for the extramathematical interpretation of Δ, it is not essential for my model, where Δ can be freely replaced with its closed convex hull $\text{cl}(\text{co}\,\Delta)$ without changing the utility representations or the behavioral axioms.

3 Model

Let a binary relation $\succeq$ be the decision maker's *weak preference* over $\mathcal{H}$. Note that the information set Δ is not a component of objects of choice, and the preference $\succeq$ is defined over acts rather than over pairs of acts and information.

Axiom 1 (Order)**.** *$\succeq$ is complete and transitive.*

Axiom 2 (Continuity)**.** *For any acts $f, g, h \in \mathcal{H}$, the sets*

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

are closed in $[0,1]$.

Given any act $f \in \mathcal{H}$, let

$$L(f, \Delta) = \{l \in \mathcal{L} : f(q) \succeq l \quad \text{for all } q \in \Delta\}.$$

Then every lottery $l \in L(f, \Delta)$ is inferior to all lotteries $f(q)$ induced by f via probabilistic scenarios $q \in \Delta$, including the objective probability law $p_0 \in \Delta$. This objective dominance motivates

Axiom 3 (Δ-Monotonicity)**.** *$f \succeq l$ for all acts $f \in \mathcal{H}$ and lotteries $l \in L(f, \Delta)$.*

Note that Δ-Monotonicity implies Δ'-Monotonicity for any $\Delta \subset \Delta' \subset \mathcal{P}$.

Recall that the standard Independence axiom requires

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

for all acts $f, g, h \in \mathcal{H}$ and $\alpha \in [0,1]$. Similarly to the Ellsberg Paradox, the decision maker may violate this separability if she perceives the mixture $\alpha g + (1-\alpha)h$ to "hedge" ambiguity better than $\alpha f + (1-\alpha)h$. The information set Δ can provide a formal meaning to such hedging.

Given any two acts $f, g \in \mathcal{H}$, say that f is *more secure* or *less secure* than g if $L(f, \Delta) \supset L(g, \Delta)$ or $L(f, \Delta) \subset L(g, \Delta)$ respectively.

Axiom 4 (Δ-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 $\alpha g + (1-\alpha)l$,*

$$f \succeq g \quad \Rightarrow \quad \alpha f + (1-\alpha)h \succeq \alpha g + (1-\alpha)h. \tag{3.1}$$

If $h = l$, then Δ-Independence becomes Gilboa–Schmeidler's C-Independence

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

More broadly, it assumes that the preference $\alpha f + (1-\alpha)l \succeq \alpha g + (1-\alpha)l$ should be preserved 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 $\alpha g + (1-\alpha)l$.

To motivate this separability, let $u \in \mathcal{U}$ be an expected utility representation for the ranking of lotteries $l \in \mathcal{L}$. The existence of u is guaranteed by Order, Continuity, and C-Independence. For any act $f \in \mathcal{H}$, define its *security level*

$$M(f) = \min_{q \in \mathrm{cl}(\Delta)} u(f(q))$$

to be the expected utility of f under the least favorable probabilistic scenario in the closure of Δ. Note that f is more secure than g if and only if $M(f) \geq M(g)$.

Define the *security premium*[3] of any mixture $\alpha f + (1-\alpha)g$ as

$$SP(\alpha, f, g) = M(\alpha f + (1-\alpha)g) - [\alpha M(f) + (1-\alpha)M(g)] \geq 0.$$

This premium is non-negative because V is concave. It will be shown below (Lemma 6) that for all $\alpha \in [0,1]$ and acts $f, g, h \in \mathcal{H}$,

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

if and only if there exists $l \in \mathcal{L}$ 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 $\alpha g + (1-\alpha)l$. In this case, the mixture $\alpha f + (1-\alpha)h$ can be interpreted to *hedge* ambiguity better than $\alpha g + (1-\alpha)h$ because the security premium $SP(\alpha, f, h)$ is greater or equal than that of $SP(\alpha, g, h)$. In this interpretation, ambiguity aversion should motivate separability (3.1) in Δ-Independence.

Note that for all mixtures $\alpha f + (1-\alpha)g$, Uncertainty Aversion

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

is also implied by Δ-Independence (together with Order and Continuity) because $SP(\alpha, f, g) \geq 0 = SP(\alpha, g, g)$.

Say that $\succeq$ is *extremely cautious* if $f(q) \succeq f$ for all $f \in \mathcal{H}$ and $q \in \Delta$. The extremely cautious decision maker can choose an ambiguous act f over a constant lottery $l \in \mathcal{L}$ only if $f(q) \succeq l$ for all $q \in \Delta$. Otherwise, she should avoid ambiguity and choose l over f because $l \succ f(q) \succeq f$ for some $q \in \Delta$.

Theorem 1. *Axioms 1–4 hold if and only if $\succeq$ is represented by*

$$U(f) = (1-\varepsilon)u(f(p)) + \varepsilon \min_{q \in \mathrm{cl}(\Delta)} u(f(q)), \tag{3.2}$$

where $\varepsilon \in [0,1]$, $p \in \mathrm{cl}(\mathrm{co}\,\Delta)$, and $u \in \mathcal{U}$.

Moreover, if $\succeq$ is not extremely cautious, then

[3] This definition is analogous to the notions of security premium in Kopylov [26] and risk premium for monetary gambles in Kreps [28, p.74].

(i) it has another representation (3.2) *with components* $\varepsilon' \in [0,1]$, $p' \in \mathcal{P}$, *and* $u' \in \mathcal{U}$, *if and only if* $\varepsilon' = \varepsilon$, $p' = p$, *and* $u' = \alpha u + \beta$ *for some* $\alpha > 0$ *and* $\beta \in \mathbb{R}$;

(ii) for all $f, g \in \mathcal{H}$ *such that* f *is more secure than* g,

$$f(p) \succ g(p) \quad \Rightarrow \quad f \succ g, \tag{3.3}$$

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

If $\varepsilon = 0$ and Δ is not singleton, then the decision maker does not know objective probabilities precisely, but still maximizes expected utility

$$U(f) = u(f(p))$$

with respect to her subjective belief $p \in \mathrm{cl}(\mathrm{co}\,\Delta)$.

In general, representation (3.2) evaluates every act f via the ε-mixture of the least favorable belief in the set $\mathrm{cl}(\Delta)$ and the belief $p \in \mathrm{cl}(\mathrm{co}\,\Delta)$ that is common for all acts.[4] This representation can be written in the MEU form

$$U(f) = \min_{q \in \Pi} u\left(f(q)\right), \tag{3.4}$$

where the set of priors

$$\Pi = (1 - \varepsilon)\{p\} + \varepsilon \mathrm{cl}(\mathrm{co}\,\Delta) \tag{3.5}$$

is the *epsilon contamination* of the probability measure $p \in \mathrm{cl}(\mathrm{co}\,\Delta)$ by the closed convex hull of the information set Δ. Therefore, Theorem 1 refines the multiple priors model and delivers the added structure (3.5) for the set Π. Note that the exogenous information set Δ is necessary for this refinement. If Δ can be picked arbitrarily, then any Π has structure (3.5), most trivially for $\varepsilon = 1$ and $\Delta = \Pi$.

The subjective belief p manifests itself directly through the rankings (3.3) 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 $\mathcal{P}$ such that the comparison of the induced lotteries $f(p) \succ (\succeq) g(p)$ represents all the rankings $f \succ (\succeq) g$ when f is more secure than g.

In contrast with the standard models of subjective probability, the use of the belief p is restricted in a way that allows the Ellsberg-type ambiguity aversion. The rankings $f(p) \succ g(p)$ and $g \succ f$ may coexist whenever g is more secure than f. In particular, the decision maker may prefer to bet on event A rather

[4]Equivalently, p can be mixed with the least favorable belief for the act f in $\mathrm{cl}(\mathrm{co}\,\Delta)$ because

$$\min_{q \in \mathrm{cl}(\Delta)} u(f(q)) = \min_{q \in \mathrm{cl}(\mathrm{co}\,\Delta)} u(f(q)).$$

than B even if she believes that A is more likely than B, but finds the bet on B more secure. Intuitively, she may exhibit these preference reversals because she is not fully confident in her assessment of probabilities p. The parameter $1-\varepsilon$ in representation (3.2) is then interpretable as the subjective degree of *confidence* in the belief p. Formally, this parameter can be associated with several behavioral patterns that I discuss next.

3.1 Confidence and Unambiguous Preferences

For any preference $\succeq$, define its unambiguous part $\succeq^*$ over acts $f, g \in \mathcal{H}$ via

$$f \succeq^* g \quad \Leftrightarrow \quad \alpha f + (1-\alpha)h \succeq \alpha g + (1-\alpha)h \quad \text{for all } \alpha \in [0,1] \text{ and } h \in \mathcal{H}.$$

This definition is proposed by GMM to distinguish comparisons that hold for all probabilistic scenarios in the subjective set of priors Π. According to my representation (3.2), the unambiguous preference $f \succeq^* g$ may hold even if it is objectively possible that $g(p_0) \succ f(p_0)$, but the decision maker is sufficiently confident in her subjective belief p to overcome such ambiguity concerns.

Consider two preferences $\succeq_1$ and $\succeq_2$ with common non-singleton information set $\Delta \subset \mathcal{P}$.

Theorem 2. *If preferences $\succeq_1$ and $\succeq_2$ satisfy Axioms 1–4, then the following statements are equivalent*

(i) for all $f, g \in \mathcal{H}$,

$$f \succeq_1^* g \quad \Rightarrow \quad f \succeq_2^* g, \tag{3.6}$$

(ii) for all acts $f \in \mathcal{H}$ and lotteries $l \in \mathcal{L}$,

$$f \succeq_1 l \quad \Rightarrow \quad f \succeq_2 l, \tag{3.7}$$

(iii) $\succeq_1$ and $\succeq_2$ have representations (3.2) *with tuples* (u, p_1, ε_1) *and* (u, p_2, ε_2) *such that $\varepsilon_2 \leq \varepsilon_1 = 1$, or $\varepsilon_2 \leq \varepsilon_1 < 1$ and $p_2 \in \frac{\varepsilon_1-\varepsilon_2}{1-\varepsilon_2}\mathrm{cl}(\mathrm{co}\,\Delta) + \frac{1-\varepsilon_1}{1-\varepsilon_2}p_1$.*

Condition (3.7) is the comparative definition of ambiguity aversion used by Epstein [11] and Ghirardato and Marinacci [19]. Either condition (3.6) or (3.7) is equivalent to the inclusion $\Pi_2 \subset \Pi_1$ for the sets of priors Π_i in the multiple priors representations (3.4) for the preferences $\succeq_i$. Statement (iii) rewrites this inclusion for the ε-contamination case when $\Pi_1 = \varepsilon_1\Delta + (1-\varepsilon_1)\{p_1\}$ and $\Pi_2 = \varepsilon_2\Delta + (1-\varepsilon_2)\{p_2\}$. The inequality $\varepsilon_2 \leq \varepsilon_1$ is derived here together with the requirement that the subjective beliefs p_1 and p_2 diverge within a limited range determined by ε_1, ε_2 and Δ. Therefore, each of the behavioral conditions (3.6) and (3.7) is sufficient, but not necessary for $\varepsilon_2 \leq \varepsilon_1$.

To make the characterization more precise, assume that Δ is not an interval, that is, Δ contains at least three linearly independent scenarios $q, q', p'' \in \mathcal{P}$

such that any vanishing linear combination $\alpha q + \alpha' q' + \alpha'' q'' = 0$ must have zero coefficients $\alpha = \alpha' = \alpha'' = 0$. For any act $f \in \mathcal{H}$, let

$$R_i(f) = \{l \in \mathcal{L} : f \succeq_i^* l \ \text{ or } l \succeq_i^* f\}.$$

be the set of all lotteries l such that the choice between f or l is unambiguous according to the ranking $\succeq_i^*$.

Theorem 3. *Suppose that Δ is not an interval, and preferences $\succeq_1$ and $\succeq_2$ satisfy Axioms 1–4. Then there exists an act $f \in \mathcal{H}$ such that*

$$R_1(f) \subsetneq R_2(f) \tag{3.8}$$

if and only if $\succeq_1$ and $\succeq_2$ have utility representations (3.2) *with $u_1 = u_2$ and $\varepsilon_1 > \varepsilon_2$.*

This result does not impose any constraints on the beliefs p_1 and p_2 except for the regular inclusion $p_1, p_2 \in \mathrm{cl}(\mathrm{co}\,\Delta)$. The comparison $\varepsilon_1 > \varepsilon_2$ is characterized in terms of unambiguous preferences between acts f and lotteries l. If $R_1(f)$ is a strict subset of $R_2(f)$ for some $f \in \mathcal{F}$, then $\succeq_2$ allows to compare f unambiguously with more lotteries than $\succeq_1$. This suggests that there is more subjective confidence in the belief p_2 than in p_1, which translates into the strict inequality $1 - \varepsilon_2 > 1 - \varepsilon_1$.

To characterize the inequality $\varepsilon_1 \leq \varepsilon_2$, one can negate (3.8) and require that for all $f \in \mathcal{H}$, $R_1(f)$ is not a proper subset of $R_2(f)$.

4 Bayesian Updating

Upon observing an event $E \in \Sigma$, the decision maker should update both the information set Δ and her subjective belief p in representation (3.2). To make the Bayesian rule well-defined for both Δ and p, assume that $lEl' \succ l'$ for some lotteries $l, l' \in \mathcal{L}$. Call such events *non-null.*

Accordingly, say that E is *null* if $l' \succeq lEl'$ for all $l, l' \in \mathcal{L}$. Even if E is null, the strict preference $l' \succ lEl'$ can still hold for representation (3.2) with $p(E) = 0$ and $\max_{q \in \Delta} q(E) > 0$. Yet p cannot be updated via the Bayesian rule in this case.

For every $q \in \mathcal{P}$ such that $q(E) > 0$, let $q|E$ be the Bayesian update of q conditional on E, that is,

$$(q|E)(A) = \frac{q(A \cap E)}{q(E)}$$

for all events $A \in \Sigma$.

Let $\succeq_E$ be the decision maker's weak preference over acts in $\mathcal{H}$ after she observes the event E. Then she should conclude that the objective probability law $p_0 \in \Delta$ satisfies $p_0(E) > 0$, and $p_0|E$ belongs to the set

$$\Delta|E = \{q|E : q \in \Delta, q(E) > 0\}.$$

If this set is empty, then $q(E) = 0$ for all $q \in \Delta$, and E is null by Δ-Monotonicity.

For any $f \in \mathcal{H}$, let

$$L(f, \Delta|E) = \{l \in \mathcal{L} : f(q) \succeq l \quad \text{for all } q \in \Delta|E\}.$$

If $L(f, \Delta|E) \supset L(g, \Delta|E)$, say that the act $f \in \mathcal{H}$ is *more secure* than $g \in \mathcal{H}$ on the event E.

Impose Order, Continuity, $\Delta|E$-Monotonicity,[5] and $\Delta|E$-Independence on the preference $\succeq_E$. Under the conditions of Theorem 1, $\succeq$ and $\succeq_E$ reveal subjective beliefs $p \in \mathrm{cl}(\mathrm{co}\,\Delta)$ and $p_E \in \mathrm{cl}(\mathrm{co}\,\Delta|E)$ respectively, but p_E need not equal $p|E$ or be the result of any other specific updating formula.

To characterize the Bayesian rule for subjective beliefs p_E, consider

Axiom 5 (Δ-Dynamic Consistency)**.** *For all $f, g, h \in \mathcal{H}$ such that fEh is less secure than gEh, but f is more secure than g on E,*

$$fEh \succeq gEh \quad \Rightarrow \quad f \succeq_E g.$$

Indeed, conditions (3.3) imply that for all $f, g, h \in \mathcal{H}$, if f is more secure than g on E, then

$$f(p_E) \succeq g(p_E) \quad \Rightarrow \quad f \succeq_E g,$$

and if fEh is less secure than gEh, then

$$fEh \succeq gEh \quad \Rightarrow \quad (fEh)(p) \succeq (gEh)(p) \quad \Rightarrow \quad f(p|E) \succeq g(p|E).$$

If $p_E = p|E$, then Δ-Dynamic Consistency follows.

Theorem 4. *Suppose that E is a non-null event, and preferences $\succeq$ and $\succeq_E$ satisfy Axioms 1–4 with information sets Δ and $\Delta|E$ respectively. Then Δ-Dynamic Consistency holds if and only if $\succeq$ and $\succeq_E$ are represented by*

$$\begin{aligned} U(f) &= (1-\varepsilon)u\left(f(p)\right) + \varepsilon \min_{q \in \mathrm{cl}(\Delta)} u\left(f(q)\right) \\ U_E(f) &= (1-\lambda)u\left(f(p|E)\right) + \lambda \min_{q \in \mathrm{cl}(\Delta|E)} u\left(f(q)\right) \end{aligned} \tag{4.1}$$

where $\varepsilon, \lambda \in [0,1]$, $u \in \mathcal{U}$, and $p \in \mathrm{cl}(\mathrm{co}\,\Delta)$ is such that $p(E) > 0$.

Moreover, if $\succeq$ is not extremely cautious, then representations (4.1) *are unique up to a positive linear transformation of u.*

In representations (4.1), both the information set Δ and the subjective belief p are updated via the Bayesian rule. This rule is imposed exogenously on the information set Δ, but the subjective belief p and its updates are derived endogenously from the observable preferences $\succeq$ and $\succeq_E$.

[5] Note that $\Delta|E$-Monotonicity implies *consequentialism* $f \sim_E fEh$ for all acts $f, h \in \mathcal{H}$ because $q(E) = 1$ and $f(q) = (fEh)(q)$ for all $q \in \Delta|E$. Thus $\succeq_E$ can be viewed as a preference over contingent acts that map E into $\mathcal{L}$.

In general, the parameters λ and ε are not related to each other, which suggests that the decision maker's confidence in her subjective belief can change arbitrarily with the arrival of new information. Unfortunately, Theorems 2 and 3 cannot be applied directly to compare the parameters λ and ε for preferences $\succeq_1=\succeq$ and $\succeq_2=\succeq_E$ because the corresponding information sets Δ and $\Delta|E$ are distinct. Moreover, despite the tight mathematical connection, the geometric sizes of these sets do not have any stable proportion.[6] It makes it impossible to compare the confidence parameters by comparing the sets of priors $\varepsilon\Delta+(1-\varepsilon)p$ and $\lambda(\Delta|E)+(1-\lambda)p|E$ for general Δ.

Yet one can still use Theorems 2 and 3 via the following construction. Let $S = S_1 \times S_2$, and every state $s = (s_1, s_2)$ be determined by two objectively unrelated sources of uncertainty. Suppose that $\Delta = \Delta_1 \times \Delta_2$, so that each $q \in \Delta$ is the Fubini product of two measures $q_1 \in \Delta_1$ on S_1 and $q_2 \in \Delta_2$ on S_2. Suppose that there is some ambiguity on S_2 so that Δ_2 is not a singleton.

Let $\mathcal{H}_2$ be the set of all acts measurable with respect to the algebra of events Σ_2 on S_2. Then for any non-null event $E = E_1 \times S_2$ that is expressed exclusively in terms of the first component $s_1 \in S_1$, the resolution of the event E preserves any $q \in \Delta$ on S_2. Thus one can compare ε and λ by applying Theorem 2 or 3 to the rankings $\succeq$ and $\succeq_E$ restricted to $\mathcal{H}_2$ because the sets $\Delta|E$ and Δ coincide when restricted to S_2.

5 Discussion

5.1 Non-Bayesian Updating of the Information Sets

My model allows to combine the Bayesian updating of subjective beliefs and non-Bayesian procedures for some information sets Δ and events E.

Say that the event E is *surprising* if $\inf_{q\in\Delta} q(E) = 0$, that is, probability scenarios in Δ allow E to have an arbitrarily small probability. This event is called surprising because it occurs with arbitrarily small probability for some scenarios in Δ. It is plausible that the information set Δ is not updated on the event E via the prior-by-prior Bayesian rule because some scenarios $q \in \Delta$ and their updates $q|E$ can be viewed as impossible if E occurs.

Let $\Gamma \subset \mathcal{P}$ be the exogenous updated information set on the event E. For

[6] For example, take $S = \{1, 2, 3, 4\}$ and $\Delta = \{q, q'\}$ such that

$$q = \left(\alpha^2, (1-\alpha)\alpha, 1-\alpha, 0\right) \quad \text{and} \quad q' = ((1-\alpha)\alpha, \alpha^2, 1-\alpha, 0)$$

for some small $\alpha > 0$. Then $\Delta|\{1, 4\}$ consists of just one point $(1, 0, 0, 0)$, but $\Delta|\{1, 2\}$ consists of two points $(\alpha, 1-\alpha, 0, 0)$ and $(1-\alpha, \alpha, 0, 0)$. Thus conditioning on the event $E = \{1, 4\}$ reduces the geometric size of Δ to zero, but conditioning on $E' = \{1, 2\}$ expands the size of Δ from an arbitrarily small number $\sqrt{2}(\alpha - 2\alpha^2)$ to $\sqrt{2}(1 - 2\alpha)$ that is close to $\sqrt{2}$ for small α.

example,

$$\Gamma = \{q|E : q \in \Delta, q(E) > \theta\}$$

for some threshold $\theta > 0$.

Proposition 5. *Suppose that E is a non-null surprising event, and preferences $\succeq$ and $\succeq_E$ satisfy Axioms 1–4 with information sets Δ and its update Γ respectively. Then Δ-Dynamic Consistency holds if and only if $\succeq$ and $\succeq_E$ are represented by*

$$\begin{aligned} U(f) &= (1-\varepsilon)u\left(f(p)\right) + \varepsilon \min_{q\in \mathrm{cl}(\mathrm{co}\,()\Delta)} u\left(f(q)\right) \\ U_E(f) &= (1-\lambda)u\left(f(p|E)\right) + \lambda \min_{q\in \mathrm{cl}(\mathrm{co}\,()\Gamma)} u\left(f(q)\right) \end{aligned} \tag{5.1}$$

where $\varepsilon, \lambda \in [0,1]$, $u \in \mathcal{U}$, $p \in \mathrm{cl}(\mathrm{co}\,\Delta)$ is such that $p(E) > 0$, and $p|E \in \mathrm{cl}(\mathrm{co}\,\Gamma)$.

Moreover, if $\succeq$ and $\succeq_E$ are not extremely cautious, then representations (5.1) are unique up to a positive linear transformation or u.

Representation (5.1) combines the Bayesian update $p|E$ of the subjective belief p with an arbitrary exogenous updated information set Γ such that $p|E \in \mathrm{cl}(\mathrm{co}\,\Gamma)$. If this inclusion does not hold, then Δ-Dynamic Consistency should be violated as well. However, the freedom in choosing the update Γ may be an embarrassment of riches. Obtaining some endogenous structure for Γ is another possible research problem.

5.2 Elicitation of Beliefs and Parameter ε

In general, one can elicit p and ε from the preferences $\succeq$ and the information set Δ 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

$$\begin{aligned} \pi(A) &= \max\{\gamma \in [0,1] : xAy \succeq l_\gamma\} = \min_{q\in\Pi} q(A) \\ \pi_*(A) &= \max\{\gamma \in [0,1] : l_\gamma \in L(xAy, \Delta)\} = \min_{q\in \mathrm{cl}(\Delta)} q(A). \end{aligned}$$

Here $\pi(A)$ measures the decision maker's willingness to bet on the event A and $\pi_*(A)$ specifies her security level for this bet.

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

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

(ii) There is $A \in \mathcal{A}$ 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 $A \in \mathcal{A}$ such that $\pi_*(A) + \pi_*(A^c) < \pi(A) + \pi(A^c) < 1$. By (3.5),

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

By summing the two equations, obtain

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

For all events $B \in \mathcal{A}$, 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.

A APPENDIX: PROOFS

Proof of Theorem 1.

Fix a non-empty set $\Delta \subset \mathcal{P}$. Suppose that $\succeq$ satisfies Order, Continuity, Δ-Monotonicity, and Δ-Independence. By Herstein–Milnor's Theorem, the ranking of lotteries has an expected utility representation $u \in \mathcal{U}$ that is unique up to a positive linear transformation. If u is constant, then (3.2) is trivial. Hereafter, assume that u is non-constant. Without loss of generality, the range of u contains the interval $I = [-1, 1]$. For any act $f \in \mathcal{H}$, let $u(f) \in \mathbb{R}^S$ be the composition of u and f. Then for any $a \in I^S$, there exists f such that $a = u(f)$.

For any $f \in \mathcal{H}$, let

$$M(f) = \inf_{q\in\Delta} u(f(q)) = \min_{q\in \mathrm{cl}(\mathrm{co}\, D)} u(f(q)).$$

Note that $l \in L(f, \Delta)$ if and only $M(f) \geq u(l)$. Thus f is more secure than g if and only if $M(f) \geq M(g)$.

Lemma 6. *For all $\alpha \in [0, 1]$ and $f, g, h \in \mathcal{H}$, $SP(\alpha, f, h) \geq SP(\alpha, g, h)$ if and only if there is $l \in \mathcal{L}$ 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. For all $\alpha \in [0, 1]$, $f, g, h \in \mathcal{H}$, and $l \in \mathcal{L}$,

$$M(\alpha f + (1-\alpha)l) = \alpha M(f) + (1-\alpha)u(l)$$
$$M(\alpha g + (1-\alpha)l) = \alpha M(g) + (1-\alpha)u(l).$$

By definition of security premia,

$$SP(\alpha, f, h) - SP(\alpha, g, h) = [M(\alpha f + (1-\alpha)h) - M(\alpha f + (1-\alpha)l)] + [M(\alpha g + (1-\alpha)l) - M(\alpha g + (1-\alpha)h)].$$

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 $l \in \mathcal{L}$ such that $M(\alpha f + (1-\alpha)h) = M(\alpha f + (1-\alpha)l)$. Then $\alpha f + (1-\alpha)h$ is more secure than $\alpha f + (1-\alpha)l$ and

$$M(\alpha g + (1-\alpha)l) - M(\alpha g + (1-\alpha)h) = SP(\alpha, f, h) - SP(\alpha, g, h) \geq 0,$$

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

Assume wlog that Δ is convex and closed. Otherwise, replace Δ by its closed convex hull. The function M and all the proofs below are unaffected.

The preference $\succeq$ satisfies all conditions in Theorem 1 in Gilboa-Schmeidler [21]. In particular, for all $\alpha \in [0,1]$ and $f, g \in \mathcal{H}$,

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

because M is concave. By Lemma 6 and Δ-Independence, $\succeq$ satisfies Uncertainty Aversion. Thus, there is a unique set convex and closed set $\Pi \subset \mathcal{P}$ such that $\succeq$ is represented by

$$U(f) = \min_{q \in \Pi} u(f(q)). \tag{A.1}$$

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

$$V(a) = \min_{q \in \Delta} q \cdot a \quad \text{and} \quad W(a) = \min_{q \in \Pi} q \cdot a. \tag{A.2}$$

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, concave, and satisfy

$$V(\alpha a + \vec{\gamma}) = \alpha V(a) + \gamma \quad \text{and} \quad W(\alpha a + \vec{\gamma}) = \alpha W(a) + \gamma \tag{A.3}$$

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

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

$$W(a) \geq V(a). \tag{A.4}$$

By (A.3), it is enough to show this claim for $a \in I^S$. Suppose that $W(a) < V(a)$ for some $a \in I^S$. Then $a = u(f)$ for some $f \in \mathcal{H}$, and there is $l \in \mathcal{L}$ such that $M(f) = V(a) > u(l) > W(a) = U(f)$. Thus $l \in L(f, \Delta)$, but $l \succ f$, which contradicts Δ-Monotonicity.

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 (A.2): for all $q \in \Delta$ such that $q \neq v(a)$,

$$V(a) = v(a) \cdot a < q \cdot a. \tag{A.5}$$

Similarly, the vector $w(a) \in \Pi$ is the unique minimizer in (A.2): for all $q \in \Pi$ such that $q \neq w(a)$,

$$W(a) = w(a) \cdot a < q \cdot a. \tag{A.6}$$

It follows from (A.5) and (A.6) that $v(a)$ and $w(a)$ are extreme points in Δ and Π respectively.

Lemma 7. *For any $a, b \in \mathbb{D}$, there exists $\varepsilon \geq 0$ such that*

$$w(a) - w(b) = \varepsilon(v(a) - v(b)). \tag{A.7}$$

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

$$W(a) \geq W(b) \quad \Rightarrow \quad W(a+c) \geq W(b+c). \tag{A.8}$$

By (A.3), it is sufficient to show this claim for vectors $a, b, c \in I^S$. Take acts $f, g, h \in \mathcal{H}$ such that $u(f) = a$, $u(g) = b$, and $u(h) = c$. Take a lottery $l \in \mathcal{L}$ 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 $f \succeq g$,

$$\begin{aligned} M\left(\tfrac{f+h}{2}\right) &= V\left(u\left(\tfrac{f+h}{2}\right)\right) = V\left(\tfrac{a+c}{2}\right) \geq V\left(\tfrac{a}{2}\right) = V\left(u\left(\tfrac{f+l}{2}\right)\right) = M\left(\tfrac{f+h}{2}\right) \\ M\left(\tfrac{g+l}{2}\right) &= V\left(u\left(\tfrac{g+l}{2}\right)\right) = V\left(\tfrac{b}{2}\right) \geq V\left(\tfrac{b+c}{2}\right) = V\left(u\left(\tfrac{g+h}{2}\right)\right) = M\left(\tfrac{g+h}{2}\right). \end{aligned}$$

By Lemma 6, $\frac{f+h}{2}$ is more secure than $\frac{f+l}{2}$, but $\frac{g+l}{2}$ is less secure than $\frac{g+h}{2}$. By Δ-Independence, $\frac{f+h}{2} \succeq \frac{g+h}{2}$. Therefore

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

and by (A.3), $W(a+c) \geq W(b+c)$.

Turn to (A.7). Fix any $a, b \in \mathbb{D}$. The derivatives of the functions W and V, and hence the equality (A.7), are unaffected if the vectors a and b are replaced by $a - V(a)1^*$ and $b - (W(b) + V(a) - W(a))1^*$ respectively. Wlog 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 $v(a) \cdot 1^* = v(b) \cdot 1^* = w(a) \cdot 1^* = w(b) \cdot 1^* = 1$. If $\lambda_3 \neq 0$, then (A.7) holds for $\varepsilon = \frac{\lambda_1}{\lambda_3}$. 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 (A.8), $W(a + \delta a) \geq W(b + \delta a)$, 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 (A.6), $w(a) = w(b)$. The equality (A.7) holds then 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 is a contradiction with (A.8). □

If $W = V$, then $\succeq$ is extremely cautious, $\Pi = \Delta$, and the utility representation (3.2) holds for $\varepsilon = 1$ and any $p \in \Delta$.

Lemma 8. *If $W \neq V$, then there are unique $0 \leq \varepsilon < 1$ and $p \in \Delta$ such that*

$$W(a) = \varepsilon V(a) + (1 - \varepsilon) p \cdot a \tag{A.9}$$

for all $a \in \mathbb{R}^S$. Moreover, p is the only probability measure in $\mathcal{P}$ such that for all $f, g \in \mathcal{H}$, 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 7, there is $\varepsilon \geq 0$ such that

$$w(b) - w(c) = \varepsilon(v(b) - v(c)). \tag{A.10}$$

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

$$w(a) = \varepsilon v(a) + \hat{p}, \tag{A.11}$$

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

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

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 $v(a), v(b), v(c)$ 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 7, $w(a) \in B \cap C$. It follows that $w(a) = \varepsilon v(a) + \hat{p}$.

By (A.5), (A.6), and (A.11),

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

for all $a \in \mathbb{D}$. Rockafellar [34, 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 $a \in \mathbb{R}^S$,

$$W(a) = \varepsilon V(a) + \hat{p} \cdot a. \tag{A.12}$$

By (A.4),

$$\hat{p} \cdot a \geq (1 - \varepsilon)V(a). \tag{A.13}$$

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

$$V(a) \leq v(c) \cdot a < v(b) \cdot a$$
$$V(-a) \leq -v(b) \cdot a < v(c) \cdot a.$$

On the other hand, by (A.13), $V(a) + V(-a) \geq \frac{\hat{p}}{1-\varepsilon} \cdot a + \frac{\hat{p}}{1-\varepsilon} \cdot (-a) = 0$, which is a contradiction.

(ii) $\varepsilon = 1$. By (A.13), $W = V$, which contradicts $W \neq V$.

(iii) $\varepsilon < 1$. Let $p = \frac{\hat{p}}{1-\varepsilon}$. By (A.13), $p \cdot a \geq V(a) = \min_{q\in\Delta} q \cdot a$ for all $a \in \mathbb{R}^S$. As Δ is convex and closed, then $p \in \Delta$ by the separating hyperplane argument.

Turn to the uniqueness part. The parameter $0 \leq \varepsilon < 1$ is uniquely determined by (A.10), and $p = \frac{\hat{p}}{1-\varepsilon}$ is unique.

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

$$p' \cdot u(f) > p' \cdot u(g), \quad p \cdot u(f) \leq p \cdot u(g), \text{ and } V(u(f)) = V(u(g)). \tag{A.14}$$

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 $f, g \in \mathcal{H}$ such that $u(f) = a$ and $u(g) = b$. If $p \cdot a \neq p \cdot b$, then take $f, g \in \mathcal{H}$ such that $u(f) = (p \cdot b)a$ and $u(g) = (p \cdot a)b$.

It follows from (A.14), that for any $p' \in \mathcal{P}$ such that $p' \neq p$, there are $f, g \in \mathcal{H}$ such that $f(p') \succ g(p')$, f is more secure than g, but still $g \succeq f$. The proof of Lemma 8 is complete. $\square$

Lemma 8 delivers the required utility representation (3.2) 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 the condition (3.3).

Extend the construction of the utility representation (3.2) for an arbitrary state space S with an algebra of events Σ. For any $A \in \Sigma$, let

$$\pi(A) = \min_{q\in\Pi} q(A) \quad \text{and} \quad \pi_*(A) = \min_{q\in\Delta} q(A).$$

Consider two cases.

Case 1. $\pi(A) = \pi_*(A)$ for all events $A \in \Sigma$. Let $\varepsilon = 1$ and take any $p \in \Delta$.

Case 2. $\pi(A) > \pi_*(A)$ for some $A \in \Sigma$. Define $\varepsilon < 1$ by

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

For all events $B \in \Sigma$, take

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

In both cases, 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 $\succeq$

restricted to Σ' measurable acts is represented by (3.2) with the triple (u, ε, p). Thus, p is finitely additive on all of Σ, and $\succeq$ is represented by (3.2) with the triple (u, ε, p) on all of $\mathcal{H}$. The inclusion $p \in \mathrm{cl}(\mathrm{co}\,\Delta)$ follows from the fact that p is the limit of a net $p_{\Sigma'} \in \mathrm{cl}(\mathrm{co}\,\Delta)$ in the weak* closed set $\mathcal{P}$.

Turn to necessity. Suppose that $\succeq$ is represented by

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

for some $\varepsilon \in [0,1]$, $p \in \mathrm{cl}(\mathrm{co}\,D)$, $u \in \mathcal{U}$. Order and Continuity are obvious. Δ-Monotonicity holds because $l \in L(f, \Delta)$ implies $u(l) \leq M(f) \leq u(p)$ and hence $u(l) \leq U(f)$. To verify Δ-Independence, take $\alpha \in [0,1]$, acts $f, g, h \in \mathcal{H}$, and a lottery $l \in \mathcal{L}$ 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

$$\begin{aligned}
&a_1 = U\left(\alpha f + (1-\alpha)l\right) - U\left(\alpha g + (1-\alpha)l\right) \geq 0 \\
&a_2 = M\left(\alpha f + (1-\alpha)h\right) - M\left(\alpha f + (1-\alpha)l\right) \geq 0 \\
&a_3 = M\left(\alpha g + (1-\alpha)l\right) - M\left(\alpha g + (1-\alpha)h\right) \geq 0 \\
&U\left(\alpha f + (1-\alpha)h\right) - U\left(\alpha g + (1-\alpha)h\right) = a_1 + \varepsilon(a_2 + a_3) \geq 0
\end{aligned}$$

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

Proofs of Theorems 2 and 3

Let preferences $\succeq_1$ and $\succeq_2$ have utility representations (3.2) with tuples $(\varepsilon_1, p_1, u_1)$ and $(\varepsilon_2, p_2, u_2)$. Let $\Pi_1 = \varepsilon_1 \Delta + (1-\varepsilon_1)\{p_1\}$ and $\Pi_2 = \varepsilon_2 \Delta + (1-\varepsilon_2)\{p_2\}$. GMM [18] show that for all $f, g \in \mathcal{H}$,

$$\begin{array}{lcl}
f \succeq_1^* g & \Leftrightarrow & f(q) \succeq_1 g(q) \quad \text{for all } q \in \Pi_1 \\
f \succeq_2^* g & \Leftrightarrow & f(q) \succeq_2 g(q) \quad \text{for all } q \in \Pi_2.
\end{array} \tag{A.15}$$

Each of the conditions of Theorem 2 is equivalent to the inclusion $\Pi_1 \subset \Pi_2$. For (i), it is established by GMM [18], for (ii) by Ghirardato and Marinacci [19], and for (iii), it is shown directly via

$$p_2 \in \tfrac{\varepsilon_1 - \varepsilon_2}{1-\varepsilon_2}\mathrm{cl}(\mathrm{co}\,\Delta) + \tfrac{1-\varepsilon_1}{1-\varepsilon_2}\{p_1\} \quad \Leftrightarrow$$
$$(1-\varepsilon_2)\{p_2\} + \varepsilon_2 \mathrm{cl}(\mathrm{co}\,\Delta) \subset (1-\varepsilon_1)\{p_1\} + \varepsilon_1 \mathrm{cl}(\mathrm{co}\,\Delta).$$

The equivalence of risk attitudes u_1 and u_2 is implied immediately by (i), (ii), or (iii) as well.

Turn to Theorem 3. Suppose that Δ is not an interval, $\varepsilon_1 > \varepsilon_2$, and $u_1 = u_2 = u$. Assume wlog that the range of u includes $[-2, 2]$. For all $f \in \mathcal{H}$, let

$$M(f) = \min_{q \in \mathrm{cl}(\mathrm{co}\,\Delta)} u(f(q)) \quad \text{and} \quad M^*(f) = \max_{q \in \mathrm{cl}(\mathrm{co}\,\Delta)} u(f(q)).$$

By (A.15), for all $f \in \mathcal{H}$,

$$R_1(f) = \{l : u(l) \notin [\varepsilon_1 M(f) + (1-\varepsilon_1)u(f(p_1)), \varepsilon_1 M^*(f) + (1-\varepsilon_1)u(f(p_1))]\}$$
$$R_2(f) = \{l : u(l) \notin [\varepsilon_2 M(f) + (1-\varepsilon_2)u(f(p_2)), \varepsilon_2 M^*(f) + (1-\varepsilon_2)u(f(p_2))]\}. \tag{A.16}$$

The strict inclusion $R_1(f) \subsetneq R_2(f)$ implies that $\varepsilon_1[M^*(f) - M(f)] > \varepsilon_2[M^*(f) - M(f)]$, that is, $\varepsilon_1 > \varepsilon_2$. Consider two cases.

Case 1. $p_1 = p_2 = p$. Take $A \in \Sigma$ such that

$$\pi^*(A) = \max_{q \in \mathrm{cl}(\mathrm{co}\,\Delta)} q(A) > \min_{q \in \mathrm{cl}(\mathrm{co}\,\Delta)} q(A) = \pi_*(A).$$

Take $f \in \mathcal{H}$ such that $u(f(s)) = 1$ if $s \in A$ and $u(f(s)) = 0$ if $s \notin A$. As $\varepsilon_1 > \varepsilon_2$, and $p(A) \in [\pi_*(A), \pi^*(A)]$, then by (A.16), $R_2(f)$ is a strict subset of $R_2(f)$.

Case 2. $p_1 \neq p_2$. Take $q \in \Delta$ such that q, p_1, p_2 are linearly independent. Assume that S is finite. Take $a \in \mathbb{R}^S$ such that

$$a = (q - p_2) - \frac{(q - p_2) \cdot (p_1 - p_2)}{(p_1 - p_2) \cdot (p_1 - p_2)} (p_1 - p_2).$$

Then $a \neq 0$ and $a \cdot (p_1 - p_2) = 0$. By construction $|a(s)| \leq 2$ for all $s \in S$. Take $f \in \mathcal{H}$ such that $a = u(f)$. Then $u(f(p_1)) = u(f(p_2)) \neq u(f(q))$. It follows that $M^*(f) > M(f)$. By (A.16), $R_2(f)$ is a strict subset of $R_2(f)$.

Proofs of Theorem 4 and Proposition 5

Suppose that E is a non-null event, preferences $\succeq$ and $\succeq_E$ satisfy Axioms 1–4 with information sets Δ and $\Delta|E$ respectively, and Δ-Dynamic Consistency holds.

By Theorem 1, $\succeq$ and $\succeq_E$ have representations (3.2) with components (ε, p, u) and (λ, p_E, u) where $p \in \mathrm{cl}(\mathrm{co}\,\Delta)$ and $p_E \in \mathrm{cl}(\mathrm{co}\,\Delta|E)$ respectively. If $\succeq_E$ is extremely cautious, then p_E is arbitrary and can be taken equal to $p|E$. Suppose that $\lambda < 1$.

Assume that $p_E \neq p|E$. I claim that there are acts $f, g \in \mathcal{H}$ such that

$$p_E \cdot u(f) < p_E \cdot u(g), \quad (p|E) \cdot u(f) > (p|E) \cdot u(g), \text{ and } V_E(u(f)) = V_E(u(g)). \tag{A.17}$$

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

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

By construction, $p_E \cdot a > (p|E) \cdot a$, $(p|E) \cdot b > p_E \cdot b$, $p_E \cdot a \geq V_E(a) = 0$, and $p_E \cdot b \geq V_E(b) = 0$, If $p_E \cdot a = p_E \cdot b$, then take $f, g \in \mathcal{H}$ such that $u(g) = a$ and

$u(f) = b$ (the range of u is assumed to contain $[-1, 1]$.) If $p_E \cdot a \neq p_E \cdot b$, then take $f, g \in \mathcal{H}$ such that $u(g) = (p_E \cdot b)a$ and $u(f) = (p_E \cdot a)b$.

Take $l \in \mathcal{L}$ such that $u(l) = V_E(u(f)) = V_E(u(g))$. Then for all $q \in \mathrm{cl}(\mathrm{co}\,\Delta)$,

$$u(fEl(q)) = q(E)u(f(q|E)) + (1 - q(E))u(l) \geq q(E)V_E(f) + (1 - q(E))u(l) = u(l).$$

As $u(f(q|E))$ can be arbitrarily close to $V_E(f)$,

$$M(fEl) = \inf_{q \in \Delta} u(fEl(q)) = u(l).$$

Similarly, $M(gEl) = u(l)$.

It follows from (A.17) and (3.2) that fEl is less secure than gEl, f is more secure than g on E, $fEl \succeq gEl$ because $(p|E) \cdot u(f) > (p|E) \cdot u(g)$, but $g \succ_E f$ because $p_E \cdot u(f) < p_E \cdot u(g)$ and $\lambda < 1$.

This is a contradiction with Δ-Dynamic Consistency. Thus $p_E = p|E$.

To prove Proposition 5, proceed analogously, but take l such that $u(l) \leq \min_{s \in S} u(f(s))$ and $u(l) \leq \min_{s \in S} u(g(s))$. As E is surprising, then $\inf_{q \in \Delta} q(E) = 0$ and hence, $M(fEl) = M(gEl) = u(l)$. Therefore, $p_E \neq p|E$ contradicts Δ-Dynamic Consistency as well.

References

[1] F. Anscombe and R. Aumann. A definition of subjective probability. *Annals of Mathematical Statistics*, 34:199–205, 1963.

[2] S. W. Becker and O. Brownson. What price ambiguity? or the role of ambiguity in decision making. *The Journal of Political Economy*, 72:62–73, 1964.

[3] J. Berger. An overview of robust Bayesian analysis. *Test*, 3:5–124, 1994.

[4] J. Berger and M. Berliner. Robust Bayes and empirical analysis with epsilon contaminated priors. *The Annals of Statistics*, 14:461–486, 1986.

[5] C. Camerer and M. Weber. Recent developments in modeling preferences. *Journal of Risk and Uncertainty*, 5:325–370, 1992.

[6] G. Carlier, R. Dana, and N. Shahdi. Efficient insurance contracts under epsilon contaminated utilities. *The Geneva Papers on Risk and Insurance Theory*, 28:59–71, 2003.

[7] S. H. Chew, B. Miao, and S. Zhong. Partial ambiguity. Mimeo, The National University of Singapore, 2012.

[8] S. P. Curley and J. Yates. The center and range of the probability interval as factors affecting ambiguity preference. *Organizational Behavior and Human Decision Processes*, 36:273–287, 1985.

[9] B. de Finetti. La prévision: ses lois logiques, ses sources subjectives. *Annales de l'Institute Henri Poincare*, 7:1–68, 1937. Translated and reprinted in Kyburg and Smokler, 1964.

[10] D. Ellsberg. Risk, ambiguity, and the Savage axioms. *Quaterly Journal of Economics*, 75:643–669, 1961.

[11] L. Epstein. A definition of uncertainty aversion. *Review of Economic Studies*, 66:579–608, 1999.

[12] L. Epstein and M. Schneider. Recursive multiple-priors. *Journal of Economic Theory*, 61:1–31, 2003.

[13] L. Epstein and T. Wang. Intertemporal asset pricing and Knightian uncertainty. *Econometrica*, 62:283–382, 1994.

[14] L. Epstein and J. Zhang. Subjective probabilities on subjectively unambiguous events. *Econometrica*, 69:265–306, 2001.

[15] P. C. Fishburn. *Utility Theory for Decision Making.* Wiley, New York, 1970.

[16] T. Gajdos, T. Hayashi, J.-M. Tallon, and J.-C. Vergnaud. Attitude towards imprecise information. *Journal of Economic Theory*, 140:27–65, 2008.

[17] P. Ghirardato. Revisiting savage in a conditional world. *Economic Theory*, 20:83–92, 2002.

[18] P. Ghirardato, F. Maccheroni, and M. Marinacci. Differentiating ambiguity and ambiguity attitude. *Journal of Economic Theory*, 118:133–173, 2004.

[19] P. Ghirardato and M. Marinacci. Ambiguity mase precise: A comparative foundation. *Journal of Economic Theory*, 102:251–289, 2002.

[20] I. Gilboa, F. Maccheroni, M. Marinacci, and D. Schmeidler. Objective and subjective rationality in a multiple priors model. *Econometrica*, 78:755–770, 2010.

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

[22] Y. Halevy and V. Feltkamp. A bayesian approach to uncertainty aversion. *Econometrica*, 69:265–306, 2001.

[23] E. Hanany and P. Klibanoff. Updating preferences with multiple priors. *Theoretical Economics*, 2:261–198, 2007.

[24] J. L. Hodges and E. L. Lehmann. The use of previous experience in reaching statistical decisions. *Annals of Mathematical Statistics*, 23:396–407, 1952.

[25] P. Klibanoff, M. Marinacci, and S. Mukerji. A smooth model of decision making under ambiguity. *Econometrica*, 73:1849–1892, 2005.

[26] I. Kopylov. Choice deferral and ambiguity aversion. *Theoretical Economics*, 4:199–225, 2009.

[27] I. Kopylov. Simple axioms for countably additive subjective probabilities. *Journal of Mathematical Economics*, 46:867–876, 2010.

[28] D. M. Kreps. *Notes on the Theory of Choice*. Westview Press, Boulder and London, 1988.

[29] M. Machina and D. Schmeidler. A more robust definition of subjective probability. *Econometrica*, 57:571–58, 1992.

[30] E. Moreno and J. Cano. Robust bayesian analysis with epsilon contaminations partially known. *Journal of the Royal Statistical Society: Part B*, 53:143–155, 1991.

[31] K. Nishimura and H. Ozaki. Search and Knightian uncertainty. *Journal of Economic Theory*, 119:299–333, 2004.

[32] J. Oechssler and A. Roomets. A test of mechanical ambiguity. Discussion Paper 555, University of Heidelberg, 2014.

[33] F. Ramsey. Truth and probability. In *The Foundations of Mathematics and Other Logical Essays*. Harcourt, Brace and Co., New York, 1931. First published in 1926.

[34] R. T. Rockafellar. *Convex Analysis*. Princeton University Press, Princeton, 1970.

[35] L. J. Savage. *The Foundations of Statistics*. Dover Publications Inc., New York, second revised edition, 1972.

[36] M. Siniscalchi. Vector expected utility and attitudes toward variation. *Econometrica*, 77:801–855, 2009.