Uncertainty Averse Preferences*

S. Cerreia-Vioglio[a] F. Maccheroni[b] M. Marinacci[c] L. Montrucchio[c]

[a]Department of Economics, Columbia University

[b]Department of Decision Sciences, DONDENA and IGIER, Università Bocconi

[c]Collegio Carlo Alberto, Università di Torino

Draft of August 2009

Abstract

We study uncertainty averse preferences, that is, complete and transitive preferences that are convex and monotone. We establish a representation result, which is at the same time general and rich in structure. Many objective functions commonly used in applications are special cases of this representation.

*We thank Pierre-André Chiappori, Larry Epstein, Bart Lipman, Jawwad Noor, Paolo Siconolfi, Alex Tetenov and, especially, Peter Klibanoff, Aldo Rustichini, and Tom Sargent, for some very useful comments. Part of this research was done while some of the authors were visiting the Economics Department of Boston University and the Collegio Carlo Alberto, which they thank for their hospitality. The financial support of ERC (advanced grant BRSCDP-TEA) is gratefully acknowledged.

1 Introduction

Beginning with the seminal works of David Schmeidler, several choice models have been proposed in the past twenty years in the large literature on choice under uncertainty that deals with ambiguity, that is, with Ellsberg-type phenomena. As a result, there are now a few possible models of choice under ambiguity, each featuring some violation of the classic independence axiom, the main behavioral assumption questioned in this literature.

Our purpose in this paper is to put some order in this class of models by providing a common representation that, through its properties, allows to unify and classify them. Since a notion of minimal independence among uncertain acts is, at best, elusive both at a theoretical and empirical level, the starting point of our analysis is that this common representation has to be independence-free. That is, it must not rely on any independence condition on uncertain acts, however weak it may appear.

This leads us to consider complete and transitive preferences that are monotone and convex, without any independence requirement. Besides its unifying power, this is arguably the most fundamental class of economic preferences that model decision making under uncertainty. General equilibrium results are, for example, typically based on them, as well as the classic arbitrage arguments of finance.[1]

Transitivity and monotonicity are fundamental principles of economic rationality. The former requires that decision makers be consistent across their choices, while the latter requires that they prefer acts that deliver better outcomes in each state. Convexity reflects a basic negative attitude of decision makers toward the presence of uncertainty in their choices, an attitude arguably shared by most decision makers and modelled through a preference for hedging/randomization.[2] Finally, completeness – which requires decision makers to be able to compare any pair of uncertain acts – is a common simplifying assumption that can then be weakened in subsequent analysis.[3]

We call *uncertainty averse* the preferences that satisfy these properties, that is, the complete and transitive preferences that are monotone and convex. In the paper we establish a representation for uncertainty averse preferences which is, at the same time, general and rich in structure. Specifically, in a standard Anscombe-Aumann set up, let $\mathcal{F}$ be the set of all uncertain acts $f : S \to X$, where S is a state space and X a convex outcome space, and let Δ be the set of all probability measures on S. We show that a preference $\succsim$ is uncertainty

[1] See, e.g., Rigotti, Shannon, and Strzalecki [45] and the references therein.

[2] See the classic discussions in Debreu [15, p. 101] and Schmeidler [51].

[3] Along, for example, the lines of Bewley [6]. See also the discussion in Gilboa, Maccheroni, Marinacci, and Schmeidler [28].

averse and satisfies some suitable technical conditions if and only if there are a utility index $u : X \to \mathbb{R}$ and a quasiconvex function $G : u(X) \times \Delta \to (-\infty, \infty]$, increasing in the first variable, such that the preference functional

$$V(f) = \min_{p \in \Delta} G\left(\int u(f)\, dp, p\right) \qquad \forall f \in \mathcal{F} \tag{1}$$

represents $\succsim$. We also show that u and G are essentially unique.

In this representation decision makers consider through the term $G\left(\int u(f)\, dp, p\right)$ all possible probabilities p – i.e., all possible "models," in the macroeconomics language – and the associated expected utilities $\int u(f)\, dp$ of act f. They then summarize all these evaluations by taking their minimum. The quasiconvexity of G and the cautious attitude reflected by the minimum in (1) derive from the convexity of preferences. Their monotonicity, instead, is reflected by the monotonicity of G in its first argument.

Behaviorally, G can be interpreted as an index of uncertainty aversion, as Proposition 6 shows. In particular, higher degrees of uncertainty aversion correspond to pointwise smaller indices G. Moreover, the index G can be elicited from choice data, that is, it is behaviorally determined. In fact, we show that

$$G(t, p) = \sup_{f \in \mathcal{F}} \left\{ u(x_f) : \int u(f)\, dp \leq t \right\}$$

where x_f is the certainty equivalent of act f. As a result, once the utility function u is elicited, something that can be done by standard methods, the quantity $G(t, p)$ can be recovered from choice data by determining the certainty equivalents of the acts f such that $\int u(f)\, dp \leq t$. In this way, the preference functional (1) itself can be behaviorally (e.g., through experimental analysis) determined and tested.

A fundamental feature of the representation (1) is the presence of probabilities and expected utilities, even though no independence assumption whatsoever is made on uncertain acts. In other words, the representation (1) establishes a connection between the language of preferences and the language of probabilities and utilities, in keeping with the tradition of the representation theorems in choice under uncertainty pioneered by Savage [49] and Anscombe and Aumann [3].

This connection is a key contribution of our derivation and is made possible by the dual nature of our representation, detailed in Section 6, which generalizes the duality arguments that since Savage underlie the representation results of choice under uncertainty. To illustrates the scope and explanatory power of this key connection, in Section 4 we show that the properties of the representation (1), and in particular those of the index G, allow to fully characterize two important probabilistic features of a preference relation under uncertainty, that is, probabilistic sophistication and revealed ambiguity.

1.1 Generality and Structure

Representation (1) is at the same time general and rich in structure. Due to its generality, (1) is able to unify, as special cases, many of the choice criteria commonly used to model choices under uncertainty, even when *prima facie* they may appear unrelated. Thanks to its structure, this unification is insightful since all special cases can be regarded as the result of suitable specifications of the uncertainty aversion index G. Moreover, novel specifications can be suggested by the properties of G and their derivation can be significantly simplified by having the representation (1) at hand. For the same reason, also the derivation of known specifications can be simplified.[4]

All this can be seen in Section 5, where we illustrate the scope of the representation (1). In particular, we show how (1) provides a common framework for two general classes of preferences under ambiguity, the variational preferences studied by Maccheroni, Marinacci, and Rustichini [38] and the smooth ambiguity preferences studied by Klibanoff, Marinacci, and Mukerji [34].[5]

We first consider variational preferences. The main issue in studying a special case of (1) is to determine the appropriate form of the uncertainty aversion index G. Proposition 12 shows that variational preferences correspond to additively separable functions G. Indeed, variational preferences are characterized by

$$G(t,p) = t + c(p)$$

where $c : \Delta \to [0,\infty]$ is a convex function, and in this case (1) reduces to the variational representation

$$V(f) = \min_{p \in \Delta} \left\{ \int u(f)\, dp + c(p) \right\}. \tag{2}$$

As [38] shows, the variational representation (2) includes as special cases the multiple priors model of Gilboa and Schmeidler [29] and the multiplier preferences of Hansen and Sargent ([32], [31]), which can therefore be viewed as particular specifications of an additively separable uncertainty aversion index G.[6]

[4] For example, this is the case for the variational representation (2), whose derivation becomes easier when based on the representation (1).

[5] See Ergin and Gul [21], Nau [43], and Seo [52] for works related to [34].

[6] See Strzalecki [54] for conditions on variational preferences that characterize multiplier preferences (see also Subsection 5.2.5 below).

Smooth ambiguity preferences are represented by

$$V(f)=\phi^{-1}\left(\int_{\Delta}\phi\left(\int_{S}u(f(s))\,dp(s)\right)d\mu(p)\right) \tag{3}$$

where ϕ is a continuous and strictly increasing function and μ is a probability measure on Δ. As in standard statistical decision theory, the first-order probabilities p are interpreted as possible models that govern states' realizations, while the second-order probabilities μ are priors over such models. As discussed by [34], because of ambiguity the function ϕ may not be linear. In particular, the concavity of ϕ reflects ambiguity aversion and in this case $\succsim$ is an uncertainty averse preference.

Theorem 19 shows that smooth preferences with concave ϕ correspond to the uncertainty aversion index given by

$$G(t,p)=t+\min_{\nu\in\Gamma(p)}I_t(\nu\parallel\mu) \tag{4}$$

where $I_t(\cdot\parallel\mu)$ is a suitable statistical distance function that generalizes the classic relative entropy, and $\Gamma(p)$ is the set of all second-order probabilities ν that are absolutely continuous with respect to μ and that have p as their reduced, first-order, probability measure on S.

As a result, the uncertainty averse representation (1) of smooth preferences is

$$V(f)=\phi^{-1}\left(\int_{\Delta}\phi\left(\int_{S}u(f)\,dp\right)d\mu(p)\right)=\min_{p\in\Delta}\left(\int_{S}u(f)\,dp+\min_{\nu\in\Gamma(p)}I_{\int_S u(f)dp}(\nu\parallel\mu)\right) \tag{5}$$

or, equivalently,[7]

$$V(f)=\min_{\nu\in\Delta^{\sigma}(\mathcal{B}(\Delta),\mu)}\left\{\int_{\Delta}\left(\int_{S}u(f)\,dq\right)d\nu(q)+I_{\int_{\Delta}\left(\int_S u(f)dq\right)d\nu(q)}(\nu\parallel\mu)\right\}. \tag{6}$$

Representation (6), established in Proposition 20, is especially interesting. In fact, as we discuss in detail in Section 5.2.2, it suggests a novel prior uncertainty interpretation of the smooth model in which each prior ν is considered via its expected utility

$$\int_{S}u(f)\,dp=\int_{\Delta}\left(\int_{S}u(f)\,dq\right)d\nu(q)$$

suitably weighted through its distance $I_{\int_{\Delta}\left(\int_S u(f)dq\right)d\nu(q)}(\nu\parallel\mu)$ from the reference prior μ. The resulting term

$$\int_{\Delta}\left(\int_{S}u(f)\,dq\right)d\nu(q)+I_{\int_{\Delta}\left(\int_S u(f)dq\right)d\nu(q)}(\nu\parallel\mu)$$

is then minimized over all priors ν in $\Delta^{\sigma}(\mathcal{B}(\Delta),\mu)$, in keeping with the uncertainty averse nature of the representation.

[7] $\Delta^{\sigma}(\mathcal{B}(\Delta),\mu)$ is the set of all second-order probabilities ν that are absolutely continuous with respect to μ.

In the important exponential case $\phi(t) = -e^{-\theta t}$, Corollary 22 shows that (4) takes the form

$$G(t,p) = t + \frac{1}{\theta} \min_{\nu \in \Gamma(p)} R(\nu \parallel \mu)$$

that is, $I_t(\cdot \parallel \mu)$ reduces to the relative entropy $R(\cdot \parallel \mu)$.[8] In this case the uncertainty averse representations (5) and (6) take the form

$$\begin{aligned} V(f) &= \min_{p \in \Delta} \left\{ \int u(f)\, dp + \frac{1}{\theta} \min_{\nu \in \Gamma(p)} R(\nu \parallel \mu) \right\} \qquad (7) \\ &= \min_{\nu \in \Delta^{\sigma}(\mathcal{B}(\Delta),\mu)} \left\{ \int_{\Delta} \left(\int_S u(f)\, dq \right) d\nu(q) + \frac{1}{\theta} R(\nu \parallel \mu) \right\}. \end{aligned}$$

This preference functional is also variational, with $c(p) = \frac{1}{\theta} \min_{\nu \in \Gamma(p)} R(\nu \parallel \mu)$. The exponential case thus turns out to be both smooth and variational. Our last result on smooth preferences, Theorem 23, shows that the overlap between these two classes of preferences is indeed characterized by functions ϕ that are constant absolute risk averse (CARA), that is, that have either the form $\phi(t) = -\alpha e^{-\theta t} + \beta$ or $\phi(t) = \alpha t + \beta$, with $\alpha, \theta > 0$ and $\beta \in \mathbb{R}$.

Inter alia, all these results shed light on the relations between smooth and variational preferences by showing that, first, (1) is the general representation that encompasses them as special cases, and, second, that the CARA case can be regarded as their overlap.

Since variational preferences feature additively separable uncertainty aversion indices, a natural class of uncertainty averse preferences to consider are those characterized by multiplicatively separable uncertainty aversion indices. To further illustrate the flexibility of the representation (1), in Section 5 we carry out this exercise, which is related to the analysis of Chateauneuf and Faro [10].

1.2 Final Remarks and Organization

Our setting admits a game against Nature interpretation, where decision makers view themselves as playing a zero-sum game against (a malevolent) Nature. This is discussed in Section 7.

The analysis of this paper is static and its dynamic extension is a natural future research topic, along the lines of Epstein and Schneider [20] and Maccheroni, Marinacci, and Rustichini [39]. In this regard, it is also important to notice that Siniscalchi [53] and Hanany and

[8] Recall that

$$R(\nu \| \mu) = \begin{cases} \int_{\Delta} \frac{d\nu}{d\mu} \log\left(\frac{d\nu}{d\mu}\right) d\mu & \text{if } \nu \ll \mu \\ \infty & \text{otherwise.} \end{cases}$$

Klibanoff [30] have recently studied in depth updating rules for uncertainty averse preferences. Specifically, Hanany and Klibanoff [30] show how the representation (1) can be used to characterize dynamically consistent update rules for uncertainty averse preferences.

Finally, to derive the results of this paper we had to establish some novel duality results for monotone quasiconcave functions. This is a further contribution of this research project, developed in Cerreia-Vioglio, Maccheroni, Marinacci, and Montrucchio [7]. In Appendices A and B we report some of these results on quasiconcave functions and we adapt them to the needs of our derivation. These quasiconcave methods are altogether different from the concave duality methods that Maccheroni, Marinacci, and Rustichini [38] use in their study of variational preferences.

The rest of the paper is organized as follows. Section 2 presents some preliminary notions, needed to establish in Section 3 the main representation results, the proofs of which are sketched in Section 6 and found in Appendix C. Section 4 discusses probabilistic sophistication and revealed ambiguity for uncertainty averse preferences. Section 5 studies some special classes of uncertainty averse preferences.

2 Preliminaries

2.1 Decision Theoretic Set Up

We consider a set S of *states of the world*, an algebra Σ of subsets of S called *events*, and a set X of *consequences*. We denote by $\mathcal{F}$ the set of all the *(simple) acts*: functions $f : S \to X$ that are Σ-measurable and take on finitely many values.

Given any $x \in X$, define $x \in \mathcal{F}$ to be the constant act such that $x(s) = x$ for all $s \in S$. With the usual slight abuse of notation, we thus identify X with the subset of the constant acts in $\mathcal{F}$. If $f \in \mathcal{F}$, $x \in X$, and $A \in \Sigma$, we denote by $xAf \in \mathcal{F}$ the act yielding x if $s \in A$ and $f(s)$ if $s \notin A$.

We assume additionally that X is a convex subset of a vector space. For instance, this is the case if X is the set of all the lotteries on a set of prizes, as it happens in the classic setting of Anscombe and Aumann [3]. Using the linear structure of X we can define in the usual way, for every $f, g \in \mathcal{F}$ and $\alpha \in [0, 1]$, the act $\alpha f + (1 - \alpha)g \in \mathcal{F}$; it yields $\alpha f(s) + (1 - \alpha)g(s) \in X$ for every $s \in S$.

We model the decision maker's *preferences* on $\mathcal{F}$ by a binary relation $\succsim$. As usual, $\succ$ and $\sim$ denote respectively the asymmetric and symmetric parts of $\succsim$. For $f \in \mathcal{F}$, an element $x_f \in X$ is a *certainty equivalent* for f if $f \sim x_f$.

2.2 Mathematical Preliminaries

We denote by $B_0(\Sigma)$ the set of all real-valued Σ-measurable simple functions – so that $u(f) \in B_0(\Sigma)$ whenever $u : X \to \mathbb{R}$ and $f \in \mathcal{F}$ – and by $B(\Sigma)$ the supnorm closure of $B_0(\Sigma)$. If T is a nonsingleton interval of the real line, set $B_0(\Sigma, T) = \{\psi \in B_0(\Sigma) : \psi(s) \in T \text{ for all } s \in S\}$.

As well known, the dual space of $B_0(\Sigma)$ (or equivalently of $B(\Sigma)$) can be identified with the set $ba(\Sigma)$ of all bounded finitely additive measures on (S, Σ). The set of probabilities in $ba(\Sigma)$ is denoted by Δ and is a weak* compact and convex subset of $ba(\Sigma)$. Elements of Δ are denoted by p or q. Finally, we denote by $\mathcal{B}(\Delta)$ the Borel σ-algebra generated by the weak* topology on Δ.

When Σ is a σ-algebra we denote by Δ^σ the set of all countably additive probabilities in Δ. In particular, given $q \in \Delta^\sigma$, we denote by $\Delta^\sigma(q)$ the set of all probabilities in Δ^σ that are absolutely continuous with respect to q; i.e., $\Delta^\sigma(q) = \{p \in \Delta^\sigma : p \ll q\}$.

Functions of the form $G : T \times \Delta \to (-\infty, \infty]$, where T is an interval of the real line, will play a key role in the paper. We denote by $\mathcal{G}(T \times \Delta)$ the class of these functions such that:

(i) G is quasiconvex on $T \times \Delta$,

(ii) $G(\cdot, p)$ is increasing for all $p \in \Delta$,

(iii) $\inf_{p \in \Delta} G(t, p) = t$ for all $t \in T$.

We denote by $\mathcal{H}(T \times \Delta)$ the class of functions in $\mathcal{G}(T \times \Delta)$ such that:

(iv) G is lower semicontinuous on $T \times \Delta$,

(v) $G(\cdot, p)$ is extended-valued continuous on T for each $p \in \Delta$.[9]

Set $\mathrm{dom} G(\cdot, p) = \{t \in T : G(t, p) < \infty\}$ and $\mathrm{dom}_\Delta G = \bigcup_{t \in T} \{p \in \Delta : G(t, p) < \infty\}$. We denote by $\mathcal{E}(T \times \Delta)$ the set of functions in $\mathcal{H}(T \times \Delta)$ that satisfy the following stronger version of (v):

(vi) $G(\cdot, p)$ is extended-valued uniformly equicontinuous on T w.r.t. $p \in \Delta$.[10]

[9] That is, $\lim_{t \to t_0} G(t, p) = G(t_0, p) \in (-\infty, \infty]$ for all $t_0 \in T$ and $p \in \Delta$. For instance, $G(t, p) = \infty$ for all $t \in T$ is continuous in this sense.

[10] That is, for every $\varepsilon > 0$ there is $\delta > 0$ such that $t, t' \in T$ and $t \le t' + \delta$ imply $G(t, p) \le G(t', p) + \varepsilon$ for all $p \in \Delta$.

This amounts to say that the functions $G(\cdot, p)$ are either real valued or constant at ∞ (that is, either $\mathrm{dom}G(\cdot, p) = T$ or $\mathrm{dom}G(\cdot, p) = \emptyset$) and, when real valued, the functions $G(\cdot, p)$ are uniformly equicontinuous on T.

A function $G : T \times \Delta \to (-\infty, \infty]$ is *linearly continuous* if the map

$$\psi \mapsto \inf_{p \in \Delta} G\left(\int \psi dp, p\right)$$

from $B_0(\Sigma, T)$ to $[-\infty, \infty]$ is extended-valued continuous. Next we show that a function is linearly continuous if it belongs to $\mathcal{H}(T \times \Delta)$, something easily verified with a routine real analysis check.

Lemma 1 *If $G \in \mathcal{H}(T \times \Delta)$, then it is linearly continuous.*

3 Uncertainty Averse Preferences

3.1 Basic Axioms

Our analysis relies on the next three main behavioral assumptions on the preference $\succsim$, which formalize the requirements of completeness, transitivity, monotonicity, and convexity that we discussed in the Introduction.

Axiom A. 1 (Weak Order) *The binary relation $\succsim$ is nontrivial, complete, and transitive.*

Axiom A. 2 (Monotonicity) *If $f, g \in \mathcal{F}$ and $f(s) \succsim g(s)$ for all $s \in S$, then $f \succsim g$.*

Axiom A. 3 (Convexity) *If $f, g \in \mathcal{F}$ and $\alpha \in (0, 1)$, $f \sim g$ implies $\alpha f + (1 - \alpha) g \succsim f$.*

These classic axioms are all falsifiable through choice behavior. In Axiom A.1, nontriviality means that $f \succ g$ for some $f, g \in \mathcal{F}$. Axiom A.2 is a monotonicity assumption, which requires that an act is preferred if, state by state, delivers a preferred outcome. Axiom A.3 is the Uncertainty Aversion Axiom of Schmeidler [51], whose interpretation in terms of negative attitude toward the presence of uncertainty dates back to Debreu [15].

Definition 2 *A binary relation $\succsim$ is an* uncertainty averse preference *if it satisfies A.1-A.3.*

As argued in the Introduction, uncertainty averse preferences are the most fundamental class of preferences that model decision making under uncertainty and exhibit a negative attitude toward uncertainty.

The next assumption is peculiar to the Anscombe-Aumann setting and imposes a standard independence axiom on constant acts, that is, acts that only involve risk and no state uncertainty.

Axiom A. 4 (Risk Independence) *If $x, y, z \in X$ and $\alpha \in (0,1)$,*

$$x \sim y \quad \Longrightarrow \quad \alpha x + (1-\alpha) z \sim \alpha y + (1-\alpha) z.$$

We now introduce some technical assumptions, which make possible the mathematical derivation in our very general set up.

Axiom A. 5 (Continuity) *If $f, g, h \in \mathcal{F}$, the sets $\{\alpha \in [0,1] : \alpha f + (1-\alpha) g \succsim h\}$ and $\{\alpha \in [0,1] : h \succsim \alpha f + (1-\alpha) g\}$ are closed.*

Axiom A.5 is a standard continuity assumption, which along with Axioms A.1 and A.2 implies the existence of a certainty equivalent x_f for each act $f \in \mathcal{F}$ (see, e.g., [38, p. 1478]).

The next assumption requires that there are arbitrarily good and arbitrarily bad outcomes. In the representation this implies that the utility function $u : X \to \mathbb{R}$ is onto; i.e., $u(X) = \mathbb{R}$.

Axiom A. 6 (Unboundedness) *For every $x \succ y$ in X, there are $z, z' \in X$ such that*

$$\frac{1}{2} z + \frac{1}{2} y \succsim x \succ y \succsim \frac{1}{2} x + \frac{1}{2} z'.$$

For some results we use an additional continuity condition.

Axiom A. 7 (Uniform Continuity) *For every $z' \prec z$ in X, there are $y' \prec y$ in X such that*

$$f, g \in \mathcal{F} \text{ and } \frac{1}{2} f(s) + \frac{1}{2} y' \precsim \frac{1}{2} g(s) + \frac{1}{2} y \quad \forall s \in S \quad \Longrightarrow \quad \frac{1}{2} x_f + \frac{1}{2} z' \precsim \frac{1}{2} x_g + \frac{1}{2} z. \tag{8}$$

Together, Axioms A.5 and A.7 form a uniform continuity condition. Axiom A.5 implies A.7 under minimal independence assumptions on acts and for this reason it is normally enough to assume A.5 in derivations that maintain some form of independence.

We close with a standard monotone continuity condition, due to Arrow [4], which will ensure in our representation results that only countably additive probabilities matter. In applications this is a convenient property because countably additive probabilities are much better behaved than probabilities that are merely finitely additive (see [11] and [38] for more on this).

Axiom A. 8 (Monotone Continuity) *If $f, g \in \mathcal{F}$, $x \in X$, $\{E_n\}_{n \in \mathbb{N}} \in \Sigma$ with $E_1 \supseteq E_2 \supseteq$ ... and $\bigcap_{n \in \mathbb{N}} E_n = \emptyset$, then $f \succ g$ implies that there exists $n_0 \in \mathbb{N}$ such that $x E_{n_0} f \succ g$.*

3.2 The Representation

We now derive our general representation (1) for uncertainty averse preferences. It relies on Axioms A.1-A.5, that is, on the original axioms of Gilboa and Schmeidler (1989), with the key exception of their independence assumption on uncertain acts, here replaced by the much weaker Axiom A.4, which applies *only* to constant acts.

Theorem 3 *Let $\succsim$ be a binary relation on $\mathcal{F}$. The following conditions are equivalent:*

(i) $\succsim$ is an uncertainty averse preference and satisfies Axioms A.4 and A.5;

(ii) there exists a nonconstant affine $u : X \to \mathbb{R}$ and a linearly continuous $G : u(X) \times \Delta \to (-\infty, \infty]$ that belongs to $\mathcal{G}(u(X) \times \Delta)$ such that, for all f and g in $\mathcal{F}$,

$$f \succsim g \Longleftrightarrow \inf_{p \in \Delta} G\left(\int u(f)\, dp, p\right) \geq \inf_{p \in \Delta} G\left(\int u(g)\, dp, p\right). \tag{9}$$

The function u is cardinally unique and, given u, there is a (unique) minimal $G^\star$ in $\mathcal{G}(u(X) \times \Delta)$ satisfying (9), given by

$$G^\star(t, p) = \sup_{f \in \mathcal{F}} \left\{ u(x_f) : \int u(f)\, dp \leq t \right\} \qquad \forall (t, p) \in u(X) \times \Delta. \tag{10}$$

Moreover, $\succsim$ has no worst consequence if and only if $\inf u(X) \notin u(X)$. In this case $G^\star$ is lower semicontinuous on $u(X) \times \Delta$.[11]

Recall that x_f is a certainty equivalent of act f. Hence, thanks to (10) the function $G^\star$ in Theorem 3 can be derived from behavioral data. In fact, once the utility function u is elicited (by standard methods), the quantity

$$\sup_{f \in \mathcal{F}} \left\{ u(x_f) : \int u(f)\, dp \leq t \right\}$$

is determined by the certainty equivalents x_f of the acts such that $\int u(f)\, dp \leq t$.

As a result, Theorem 3 guarantees that, given an uncertainty averse decision maker that satisfies the Axioms A.4 and A.5, we can elicit the precise form of the representation

$$V(f) = \inf_{p \in \Delta} G^\star\left(\int u(f)\, dp, p\right) \qquad \forall f \in \mathcal{F} \tag{11}$$

of his preference, by using purely behavioral (e.g., experimental) data.

[11] Recall that $\succsim$ has no worst consequence if for each $x \in X$ there is $y \in X$ such that $x \succ y$, and that lower semicontinuity of $G^\star$ implies that $\inf_{p \in \Delta} G^\star\left(\int u(f)\, dp, p\right) = \min_{p \in \Delta} G^\star\left(\int u(f)\, dp, p\right)$ for all $f \in \mathcal{F}$.

By Theorem 3, uncertainty averse preferences $\succsim$ that satisfy Axioms A.4 and A.5 are characterized by pairs $(u, G^{\star})$, which we call *uncertainty averse representations* of $\succsim$.[12] Such pairs have the following uniqueness property.

Proposition 4 *Let (u, G) be an uncertainty averse representation of a preference $\succsim$. Then $(\bar{u}, \bar{G})$ is another uncertainty averse representation of $\succsim$ if and only if there exist $\alpha > 0$ and $\beta \in \mathbb{R}$ such that $\bar{u} = \alpha u + \beta$ and $\bar{G}(t, p) = \alpha G(\alpha^{-1}(t - \beta), p) + \beta$ for all $(t, p) \in \bar{u}(X) \times \Delta$.*

In Theorem 3 we establish the minimality, but not the uniqueness, of the index G. The next important result shows that uniqueness holds when $\succsim$ satisfies A.6, that is, when $u(X) = \mathbb{R}$.

Theorem 5 *Let $\succsim$ be an uncertainty averse preference that satisfies A.4-A.6. Then, $G^{\star}$ defined in (10) is the unique lower semicontinuous $G \in \mathcal{G}(\mathbb{R} \times \Delta)$ for which (9) holds.*

3.3 Comparative Attitudes

Based on Ghirardato and Marinacci [26], given two preferences $\succsim_1$ and $\succsim_2$, say that $\succsim_1$ is *more uncertainty averse than* $\succsim_2$ if, for all $f \in \mathcal{F}$ and $x \in X$,

$$f \succsim_1 x \quad \Longrightarrow \quad f \succsim_2 x. \tag{12}$$

In other words, $\succsim_1$ is more uncertainty averse than $\succsim_2$ if, whenever $\succsim_1$ is "bold enough" to prefer an uncertain act f over a constant outcome x, then the same is true for $\succsim_2$.

Next we show that comparative uncertainty attitudes are determined by the functions G. Here $u_1 \approx u_2$ means that there exist $\alpha > 0$ and $\beta \in \mathbb{R}$ such that $u_1 = \alpha u_2 + \beta$.

Proposition 6 *Given two preferences $\succsim_1$ and $\succsim_2$ with uncertainty averse representations (u_1, G_1) and (u_2, G_2), the following conditions are equivalent:*

(i) $\succsim_1$ is more uncertainty averse than $\succsim_2$,

(ii) $u_1 \approx u_2$ and $G_1 \leq G_2$ (provided $u_1 = u_2$).

Given that $u_1 \approx u_2$, the assumption $u_1 = u_2$ is just a common normalization of the two utility indices. Therefore, Proposition 6 says that more uncertainty averse preference

[12] In other words, denoting by $\mathcal{U}(X)$ the set of all nonconstant affine functions from X to $\mathbb{R}$, a pair $(u, G) \in \mathcal{U}(X) \times \mathcal{G}(u(X) \times \Delta)$ is an uncertainty averse representation of $\succsim$ if G is linearly continuous, and (9) and (10) hold.

relations are characterized, up to a normalization, by pointwise smaller functions G. The function G can thus be properly interpreted as an *index of uncertainty aversion.*

Notice that $G_1 \leq G_2$ implies $\mathrm{dom}_\Delta G_2 \subseteq \mathrm{dom}_\Delta G_1$. In view of what we will momentarily show in Section 4.2, this means that higher uncertainty aversion entails larger sets of priors that the decision makers deem relevant. More uncertainty averse decision makers can thus be viewed as perceiving more model uncertainty.

We close by characterizing minimal and maximal uncertainty averse indexes. Assume $u(X) = \mathbb{R}$. Since $\inf_{p\in\Delta} G(t,p) = t$, the maximally uncertainty averse index is given by $G(t,p) = t$ for all $t \in \mathbb{R}$ and all $p \in \Delta$. Therefore, the preference functional

$$V(f) = \min_{p\in\Delta} \int u(f)\,dp = \min_{s\in S} u(f(s))$$

represents preferences that are maximally uncertainty averse.

Subjective expected utility preferences are, instead, minimally uncertainty averse. In fact, suppose $\succsim$ is a subjective expected utility preference represented by $V(f) = \int u(f)\,dq$, for some $q \in \Delta$. Its uncertainty index is $G(t,p) = t + \delta_q(p)$ for all $(t,p) \in \mathbb{R}\times\Delta$, where δ_q denotes the indicator function

$$\delta_q(p) = \begin{cases} 0 & p = q \\ \infty & p \neq q. \end{cases}$$

Suppose $G' \in \mathcal{G}(\mathbb{R}\times\Delta)$ is such that $G' \geq G$. To prove that $\succsim$ is minimally uncertainty averse we need to show that $G' = G$. We have $G'(t,p) = G(t,p) = \infty$ for all $t \in \mathbb{R}$ if $p \neq q$, while $t \leq G(t,q) \leq G'(t,q)$ for all $t \in \mathbb{R}$. Then, $G'(t,q) = \min_{p\in\Delta} G'(t,p) = t$ for all $t \in \mathbb{R}$, and so $t = G(t,q) = G'(t,q)$ for all $t \in \mathbb{R}$. We conclude that $G = G'$, as desired.

3.4 More on Continuity

As we already observed, Axiom A.7 is a uniform continuity condition when added to Axiom A.5. Since Axiom A.5 implies A.7 under minimal independence assumptions on acts, Axiom A.7 is redundant in derivations that assume some form of independence (in fact, even very weak forms of independence ensure the Lipschitzianity of the representing preference functional).

In our independence-free setting, Axiom A.7 delivers an interesting version of the representation, in which the index G belongs to $\mathcal{E}(T\times\Delta)$ and thus features stronger (and easier to check) continuity properties.

Theorem 7 *Let $\succsim$ be a binary relation on $\mathcal{F}$. The following conditions are equivalent:*

(i) $\succsim$ *is uncertainty averse and satisfies Axioms A.4-A.7;*

(ii) there exist an affine $u : X \to \mathbb{R}$, *with* $u(X) = \mathbb{R}$, *and a* $G : \mathbb{R}\times\Delta \to (-\infty, \infty]$ *that belongs to* $\mathcal{E}(\mathbb{R}\times\Delta)$ *such that, for all* f *and* g *in* $\mathcal{F}$,

$$f \succsim g \Longleftrightarrow \min_{p\in\Delta} G\left(\int u(f)\,dp, p\right) \geq \min_{p\in\Delta} G\left(\int u(g)\,dp, p\right). \tag{13}$$

The function u *is cardinally unique and, given* u, *the index* G *is unique and satisfies*

$$G(t,p) = \sup_{f\in\mathcal{F}} \left\{ u(x_f) : \int u(f)\,dp = t \right\} \qquad \forall (t,p) \in \mathbb{R}\times\Delta. \tag{14}$$

If, in addition, Σ *is a* σ*-algebra, then* $\succsim$ *satisfies Axiom A.8 if and only if there is* $q \in \Delta^\sigma$ *such that* $G(\cdot, p) \equiv \infty$ *for all* $p \notin \Delta^\sigma(q)$*; in particular* Δ *can be replaced with* $\Delta^\sigma(q)$ *in (13).*

Observe that, *inter alia*, we now have an equality sign in (14), something that simplifies the elicitation of G since less acts f have to be considered. Moreover, inspection of the proof shows that A.1-A.6 suffice to have this equality sign, and that Δ can be replaced with $\Delta^\sigma(q)$ provided Σ is a σ-algebra and $\succsim$ satisfies Axiom A.8.

4 The Role of Probabilities

As we observed in the Introduction, an essential insight provided by our representation results is the connection between preferences and beliefs given by (9). To illustrate the explanatory power of this connection, in this section we show how probabilistic sophistication and revealed ambiguity can be related to the properties of G.

4.1 Probabilistic Sophistication

In this subsection we characterize uncertainty averse preferences that are probabilistically sophisticated, an important property of preferences introduced by Machina and Schmeidler [40] that some authors, notably Epstein [19], identify with ambiguity neutrality (or absence of ambiguity altogether).

In particular, Theorem 8 will show that the properties of symmetry of the uncertainty index that guarantee probabilistic sophistication in the special case of variational preferences (see [38, Th. 14]) remarkably turn out to be necessary and sufficient for probabilistic sophistication also for the much more general class of uncertainty averse preferences. This is also surprising mathematically since we are moving from the realm of convex analysis, where the theory of rearrangement invariant Banach spaces – key for the results of this section – was originally developed by Luxemburg [37] and Chong and Rice [12], to that of quasiconvex analysis.

Given a countably additive probability q on the σ-algebra Σ, a preference relation $\succsim$ is *probabilistically sophisticated* (w.r.t. q) if, for f and g in $\mathcal{F}$,

$$q(s \in S : f(s) = x) = q(s \in S : g(s) = x) \text{ for all } x \in X \Rightarrow f \sim g,$$

while $\succsim$ satisfies the stronger property of *stochastic dominance* (w.r.t. q) if, given f and g in $\mathcal{F}$,

$$q(s \in S : f(s) \precsim x) \leq q(s \in S : g(s) \precsim x) \text{ for all } x \in X \Rightarrow f \succsim g.$$

As discussed at length in Machina and Schmeidler [40], these properties capture the fact that in comparing acts the decision maker only cares about their distribution w.r.t. q. In order to characterize these properties in terms of the dual representation (9) we need to borrow some concepts from the theory of stochastic orders. Specifically, the *convex order* $\succsim_{cx}$ on $\Delta^\sigma(q)$ is defined by

$$p \succsim_{cx} p' \text{ iff } \int \psi\left(\frac{dp}{dq}\right) dq \geq \int \psi\left(\frac{dp'}{dq}\right) dq$$

for every convex function ψ on $\mathbb{R}$. Importantly, the symmetric part $\sim_{cx}$ coincides with the identical distribution of the densities w.r.t. q.

A function $J : \Delta \to (-\infty, \infty]$ is *rearrangement invariant* (w.r.t. q) if $\operatorname{dom} J \subseteq \Delta^\sigma(q)$ and, given p and p' in $\Delta^\sigma(q)$,

$$p \sim_{cx} p' \Rightarrow J(p) = J(p')$$

while it is *Shur convex* (w.r.t. q) if $\operatorname{dom} J \subseteq \Delta^\sigma(q)$ and, given p and p' in $\Delta^\sigma(q)$,

$$p \succsim_{cx} p' \Rightarrow J(p) \geq J(p'). \tag{15}$$

Finally, we say that q in Δ^σ is *adequate* if either q is non-atomic or S is finite and q is uniform.

Theorem 8 *Let $\succsim$ be an uncertainty averse preference that satisfies A.4-A.6, A.8, and let (u, G) be an uncertainty averse representation of $\succsim$. If $q \in \Delta^\sigma$ is adequate, then the following conditions are equivalent (w.r.t. q):*

(i) $\succsim$ satisfies stochastic dominance;

(ii) $\succsim$ is probabilistically sophisticated;

(iii) $G(t, \cdot)$ *is rearrangement invariant for all* $t \in \mathbb{R}$*;*

(iv) $G(t, \cdot)$ *is Shur convex for all* $t \in \mathbb{R}$*;*

Moreover, (iv) implies (i) even if q *is not adequate.*

The proof of Theorem 8 can be found in [8], which provides a general analysis of probabilistic sophistication and uncertainty aversion. Here it is important to observe the equivalence between properties (ii) and (iii), that is, between probabilistic sophistication and a symmetry property of the index G. This shows that a suitable property of the index G characterizes uncertainty averse preferences that are probabilistically sophisticated.

4.2 Revealed Ambiguity

Uncertainty averse preferences are, by definition, characterized by a preference for hedging. Thus, in general, the preference $f \succsim g$ does not imply

$$\alpha f + (1 - \alpha) h \succsim \alpha g + (1 - \alpha) h \tag{16}$$

for all acts h and for all α in $(0, 1]$, unless $\succsim$ is a subjective expected utility preference. As argued by Ghirardato, Maccheroni, and Marinacci [25], when this implication holds it means that the preference for f over g is "strong enough" to make uncertainty aversion considerations of secondary importance.[13] In this case, [25] says that f is *unambiguously preferred* to g, written $f \succsim^* g$; formally,

$$f \succsim^* g \Longleftrightarrow \alpha f + (1 - \alpha) h \succsim \alpha g + (1 - \alpha) h \qquad \forall \alpha \in (0, 1], \forall h \in \mathcal{F}.$$

Though [25] concentrates its analysis on preferences satisfying the Certainty Independence Axiom of Gilboa and Schmeidler [29], the meaning of $\succsim^*$ remains unchanged in our independence-free setting. Moreover, it is easy to check that [25, Propositions 4 and 5] hold for any preference that satisfies A.1, A.2, A.4, and A.5 (as independently observed by Ghirardato and Siniscalchi [27]). In particular, $\succsim^*$ is well defined and admits a representation à la Bewley.

Proposition 9 *Let* $\succsim$ *be an uncertainty averse preference that satisfies A.4 and A.5, and* (u, G) *be an uncertainty averse representation of* $\succsim$*. There exists a unique nonempty, closed, and convex set* C^* *in* Δ *such that*

$$f \succsim^* g \Longleftrightarrow \int u(f)\, dp \geq \int u(g)\, dp \qquad \forall p \in C^*. \tag{17}$$

[13] See also the recent [28].

The next important result shows that C^* is, up to closure, the set of priors that the decision maker deems relevant according to the representation (1), that is, the effective domain $\mathrm{dom}_\Delta G$ of the index G (these are the priors that the decision maker considers in forming the ranking $\succsim$).

Theorem 10 *Let $\succsim$ be an uncertainty averse preference that satisfies A.4-A.6, and (u, G) an uncertainty averse representation of $\succsim$. Then, $C^* = \mathrm{cl}\,(\mathrm{dom}_\Delta G)$.*

The equality $C^* = \mathrm{cl}\,(\mathrm{dom}_\Delta G)$ is an important consistency check for our analysis. For, the set C^* is derived and interpreted directly in terms of the preference $\succsim$, without any use of the representation (1). This equality is therefore conceptually remarkable and further clarifies the interpretation of G as an uncertainty index.

We conclude by providing a differential characterization of C^*. Beyond its intrinsic interest, it connects unambiguous preferences with the families of beliefs commonly used in the study of uncertainty averse preferences and their applications (see, e.g., Hanany and Klibanoff [30] and Rigotti, Shannon, and Strzalecki [45]). If $u : X \to \mathbb{R}$ and $f \in \mathcal{F}$, set

$$\pi_u(f) = \left\{ p \in \Delta : \int u(f)\, dp \geq \int u(g)\, dp \text{ implies } f \succsim g \right\}.$$

The set $\pi_u(f)$ consists of the beliefs that rationalize the preferences of the decision maker at f.[14] Mathematically, it is the normalized Greenberg-Pierskalla's superdifferential at $u(f)$ of any functional $I : B_0(\Sigma, u(X)) \to \mathbb{R}$ such that $f \succsim g \Longleftrightarrow I(u(f)) \geq I(u(g))$.

Proposition 11 *Let $\succsim$ be an uncertainty averse preference that satisfies A.4-A.6, and (u, G) be an uncertainty averse representation of $\succsim$. Then:*

(i) $\pi_u(f) = \arg\inf_{p \in \Delta} G\left(\int u(f)\, dp, p\right)$ *for all* $f \in \mathcal{F}$.

(ii) $C^* = \mathrm{cl}\left(\mathrm{co}\left(\bigcup_{f \in \mathcal{F}} \pi_u(f)\right)\right)$.

Up to closed convex closure, C^* can thus be viewed as the collection of all beliefs that "locally" rationalize the decision maker's preferences.

[14] See [45, pag. 1169] for a similar interpretation of π_u. More precisely, in their paper X is a set of monetary payoffs and the set of supporting beliefs is computed at f, while here X is generic and the set of supporting beliefs is computed at $u(f)$.

5 Special Cases

Uncertainty averse preferences are a very general class of preferences and in this section we present important special cases that can be obtained by suitably specifying the uncertainty aversion index G.

5.1 Variational Preferences

We begin with the variational preferences of Maccheroni, Marinacci, and Rustichini [38]. A pair (u, c) is a *variational representation* of a preference $\succsim$ if $u : X \to \mathbb{R}$ is an affine function and $c : \Delta \to [0, \infty]$ is a lower semicontinuous convex function, with $\min_{p\in\Delta} c(p) = 0$, such that

$$f \succsim g \Longleftrightarrow \min_{p\in\Delta}\left\{\int u(f)\,dp + c(p)\right\} \geq \min_{p\in\Delta}\left\{\int u(g)\,dp + c(p)\right\} \tag{18}$$

for all f and g in $\mathcal{F}$.

As shown by [38], a preference admits a variational representation if and only if it is an uncertainty averse preference that satisfies both Axiom A.5 and the following weak independence axiom, discussed in detail in [38].

Axiom A. 9 (Weak Certainty Independence) *If $f, g \in \mathcal{F}$, $x, y \in X$, and $\alpha \in (0, 1)$,*

$$\alpha f + (1-\alpha)x \succsim \alpha g + (1-\alpha)x \Rightarrow \alpha f + (1-\alpha)y \succsim \alpha g + (1-\alpha)y.$$

In this case, $\succsim$ is said to be a *variational preference*. A variational preference $\succsim$ satisfies A.4 (which is implied by A.9) and, setting

$$G(t, p) = t + c(p), \tag{19}$$

the pair (u, G) clearly represents $\succsim$ in the sense of (9). More is actually true:

Proposition 12 *Let $u : X \to \mathbb{R}$ be affine with $u(X) = \mathbb{R}$. If (u, c) is a variational representation of $\succsim$, then, setting*

$$G(t, p) = t + c(p) \qquad \forall (t, p) \in \mathbb{R}\times\Delta \tag{20}$$

the pair (u, G) is an (additively separable) uncertainty averse representation of $\succsim$.

Conversely, if (u, G) is an additively separable uncertainty averse representation of $\succsim$, i.e.,

$$G(t, p) = \gamma(t) + c(p) \qquad \forall (t, p) \in \mathbb{R}\times\Delta$$

for some $\gamma : \mathbb{R} \to \mathbb{R}$ and $c : \Delta \to [0, \infty]$ with $\inf_{p\in\Delta} c(p) = 0$, then γ is the identity and (u, c) is a variational representation of $\succsim$.

Variational representations are thus nothing but additively separable uncertainty averse representations. Moreover, for variational preferences $\operatorname{dom}_\Delta G = \operatorname{dom} c$, and so $C^* = \operatorname{cl}(\operatorname{dom} c)$ by Theorem 10.

5.2 Smooth Ambiguity Preferences

The smooth ambiguity preferences studied by Klibanoff, Marinacci, and Mukerji [34] provide another example of uncertainty averse preferences. In this subsection we study the properties of their uncertainty averse representation.

A triplet (u, ϕ, μ) is a *smooth (ambiguity) representation* of a preference $\succsim$ if $u : X \to \mathbb{R}$ is an affine function, $\phi : \mathbb{R} \to \mathbb{R}$ is a strictly increasing function, and μ is a countably additive Borel probability measure on Δ such that

$$f \succsim g \Longleftrightarrow \int_\Delta \phi\left(\int_S u(f(s))\,dp(s)\right) d\mu(p) \geq \int_\Delta \phi\left(\int_S u(g(s))\,dp(s)\right) d\mu(p) \tag{21}$$

for all $f, g \in \mathcal{F}$.[15]

Throughout the paper we will consider the ϕ concave case that, as mentioned in the Introduction, corresponds to the case when the smooth representation (21) is uncertainty averse. In order to establish the uncertainty averse representation of these smooth preferences, we need to introduce a family of statistical distance functions.[16]

5.2.1 A Family of Statistical Distance Functions

Denote by $\Delta^\sigma(\mathcal{B}(\Delta), \mu)$ the set of all (second-order) countably additive Borel probability measures on Δ that are absolutely continuous with respect to μ. In particular, given a $\nu \in \Delta^\sigma(\mathcal{B}(\Delta), \mu)$, denote by $d\nu/d\mu$ the Radon-Nikodym derivative of ν with respect to μ. Moreover, $\phi^* : \mathbb{R} \to [-\infty, \infty)$ is the concave conjugate of ϕ, given by $\phi^*(z) = \inf_{k \in \mathbb{R}} \{kz - \phi(k)\}$.

For all $t \in \mathbb{R}$, define $I_t(\cdot \parallel \mu) : \Delta^\sigma(\mathcal{B}(\Delta), \mu) \to [-\infty, \infty]$ by

$$I_t(\nu \parallel \mu) = \phi^{-1}\left(\inf_{k \geq 0}\left[kt - \int \phi^*\left(k\frac{d\nu}{d\mu}\right) d\mu\right]\right) - t. \tag{22}$$

The function $I_t(\cdot \parallel \mu)$ is a statistical distance on $\Delta^\sigma(\mathcal{B}(\Delta), \mu)$, as next we show.

---

[15] In richer settings (whose specification is beyond the scope of this paper), Ergin and Gul [21], Klibanoff, Marinacci, and Mukerji [34], Nau [43], and Seo [52] provide behavioral conditions that underlie the representation (21). Observe that, when needed, ϕ and ϕ^{-1} denote the extended-valued continuous extentions of ϕ and ϕ^{-1} from $[-\infty, \infty]$ to $[-\infty, \infty]$. See (63) in Appendix B.

[16] See [36] for a thorough study of statistical distance functions. Because of their ancillary nature for our analysis, for brevity we omit the proofs of the results on these functions given in Section 5.2.1. They are available upon request.

Proposition 13 *For all $t \in \mathbb{R}$,*

(i) $I_t(\mu \parallel \mu) = 0$ and $I_t(\nu \parallel \mu) \geq 0$ for all $\nu \in \Delta^\sigma(\mathcal{B}(\Delta), \mu)$;

(ii) $I_t(\cdot \parallel \mu)$ is quasiconvex;

(iii) $I_t(\cdot \parallel \mu)$ is lower semicontinuous and coercive, i.e., the lower contour sets $\{\nu \in \Delta^\sigma(\mu) : I_t(\nu \parallel \mu) \leq c\}$ are weakly compact in $\Delta^\sigma(\mathcal{B}(\Delta), \mu)$ for all $c \geq 0$.

Example 14 When ϕ is the identity, i.e., $\phi(t) = t$ for all $t \in \mathbb{R}$, then

$$I_t(\nu \parallel \mu) = \begin{cases} 0 & \text{if } \nu = \mu \\ +\infty & \text{else.} \end{cases}$$

That is, I_t has a discrete form where all $\nu \neq \mu$ are infinitely distant from μ. ▲

Example 15 The classic relative entropy $R(\nu \parallel \mu)$ is an example of function I_t. For, consider $\phi(t) = -e^{-\theta t}$, with $\theta > 0$. Simple algebra based on Proposition 18 below shows that

$$I_t(\nu \parallel \mu) = \frac{1}{\theta} R(\nu \parallel \mu) \qquad \forall t \in \mathbb{R}.$$

In particular, when $\theta = 1$ we get $I_t(\nu \parallel \mu) = R(\nu \parallel \mu)$ for all $t \in \mathbb{R}$. Notice that in this special case I_t does not depend on t. ▲

In a different context, this family of statistical distances has been considered in Mathematical Finance by Frittelli [23] and Bellini and Frittelli [5]. There is an interesting relation between the degree of concavity of ϕ and the magnitude of the induced distance I_t.

Proposition 16 *Suppose Σ is nontrivial. Then, given two strictly increasing and concave functions $\phi_1, \phi_2 : \mathbb{R} \to \mathbb{R}$, the following conditions are equivalent:*

(i) ϕ_1 is more concave than ϕ_2;[17]

(ii) $I_t^1(\cdot \parallel \mu) \leq I_t^2(\cdot \parallel \mu)$ for all $\mu \in \Delta^\sigma(\mathcal{B}(\Delta))$ and $t \in \mathbb{R}$.

In particular, $\phi_1 \approx \phi_2$ implies $I_t^1 = I_t^2$. This means, *inter alia*, that in terms of I the functions ϕ are unique up to positive linear transformations, and can therefore be normalized.

We now introduce a class of functions for which it is relatively easy to compute I_t. Here it is convenient to normalize ϕ by setting $\phi(0) = 0$ and $\phi'(0) = 1$.

[17] That is, there exists a strictly increasing and concave $h : \phi_2(\mathbb{R}) \to \mathbb{R}$ such that $\phi_1 = h \circ \phi_2$.

Definition 17 *A normalized function $\phi : \mathbb{R} \to \mathbb{R}$ is* order Orlicz *if it is strictly increasing, strictly concave, differentiable, and there exists $\alpha > 1$ such that $k\phi'(k)/\phi(k) \geq \alpha/(\alpha - 1)$ for $k < 0$ small enough and $k\phi'(k)/\phi(k) \leq \alpha/(\alpha + 1)$ for $k > 0$ large enough.*

Order Orlicz functions are thus characterized by "tail" conditions on the elasticities $k\phi'(k)/\phi(k)$ of ϕ. The normalized negative exponential is an example of order Orlicz function.

Proposition 18 *If ϕ is order Orlicz, then*

$$I_t(\nu \parallel \mu) = \phi^{-1}\left(\int (\phi \circ \psi)\left(k(\nu)\frac{d\nu}{d\mu}\right) d\mu\right) - t \qquad \forall t \in \mathbb{R}, \nu \in \mathrm{dom} I_t(\cdot \parallel \mu) \tag{23}$$

where $\psi = (\phi')^{-1}$ and $k(\nu) \in (0, \infty)$ is the only solution to the equation

$$\int \psi\left(k\frac{d\nu}{d\mu}\right) d\nu = t. \tag{24}$$

In other words, when ϕ is order Orlicz, the index I_t can be computed in two stages. First, $k(\nu)$ is determined via (24), and then it is used to determine I_t via (23). This procedure is known (see [33]), our contribution is to identify a class of functions in which it works (see also [50]).

5.2.2 Uncertainty Averse Representation

We can now state the announced representation. A piece of notation: $p = \int_\Delta q d\nu(q)$ means

$$p(A) = \int_\Delta q(A)\, d\nu(q) \qquad \forall A \in \Sigma. \tag{25}$$

Theorem 19 *Let $u : X \to \mathbb{R}$ be an affine function with $u(X) = \mathbb{R}$, $\phi : \mathbb{R} \to \mathbb{R}$ a strictly increasing and concave function, and μ a countably additive Borel probability measure on Δ. The following conditions are equivalent:*

(i) (u, ϕ, μ) is a smooth representation of $\succsim$,

(ii) (u, G) is an uncertainty averse representation of $\succsim$, where, for all $(t, p) \in \mathbb{R} \times \Delta$,

$$G(t, p) = t + \min_{\nu \in \Gamma(p)} I_t(\nu \parallel \mu) \tag{26}$$

with

$$\Gamma(p) = \left\{\nu \in \Delta^\sigma(\mathcal{B}(\Delta), \mu) : p = \int_\Delta q d\nu(q)\right\}$$

under the convention $G(\cdot, p) \equiv \infty$ when $\Gamma(p) = \emptyset$.

The important part of Theorem 19 is (26), which provides an explicit formula for the uncertainty aversion index G in the smooth case. We now discuss this formula, which leads to a novel "prior uncertainty" interpretation of the smooth representation, illustrated by Proposition 20. We first consider the key element $\Gamma(p)$.

Interpretation of $\Gamma(p)$ The term $\Gamma(p)$ has a very simple decision theoretic interpretation in terms of the standard operation of reduction of compound lotteries (i.e., of averaging of second-order probability measures, in our general setting). In fact, $\Gamma(p)$ is nothing but the set of all second-order probabilities ν that are absolutely continuous with respect to μ and that have p as their reduced, first-order, probability measure on S. This is best understood when the support of μ is finite, say $\operatorname{supp}\mu = \{q_1, ..., q_n\}$. In this case we can identify μ with a vector $(\mu_1, ..., \mu_n) \in \Delta_n$, where Δ_n denotes the simplex in $\mathbb{R}^n$. Thus, $\Delta^\sigma(\mathcal{B}(\Delta), \mu)$ can be identified with Δ_n, and so

$$\Gamma(p) = \left\{\nu \in \Delta_n : p(A) = \sum_{i=1}^{n} q_i(A)\,\nu_i \text{ for all } A \in \Sigma\right\} \qquad \forall p \in \Delta.$$

In other words, $\Gamma(p)$ is the set of all possible weights $\nu = (\nu_1, ..., \nu_n) \in \Delta_n$ such that p can be written as a convex combination of the probabilities q_i in $\operatorname{supp}\mu$. Hence, each $\nu \in \Gamma(p)$ can be regarded as a second order probability that, through the standard reduction operation (25) – i.e., $p = \sum_{i=1}^{n} q_i\nu_i$ in this finite support case – reduces to the first order probability p.

Summing up, $\Gamma(p)$ is the set of all second order probabilities associated, via the reduction operation (25), to each given first order probability. It is easy to see that to distinct p_1 and p_2 correspond disjoint sets of second order probabilities that reduce to them; that is, $\Gamma(p_1) \cap \Gamma(p_1) = \emptyset$ if $p_1 \neq p_2$.

Prior Uncertainty As mentioned in the Introduction, in the smooth model the second order probabilities are viewed as priors on the first order probabilities, which in turn describe all possible probabilistic models that stochastically determine acts' outcomes.

The priors in $\Gamma(p)$ enter formula (26) through their distance $I_t(\nu \parallel \mu)$ from the reference prior μ. In particular, the least among these distances – that is, $\min_{\nu\in\Gamma(p)} I_t(\nu \parallel \mu)$ – is considered. This minimum can be viewed as the distance according to I_t between the reference prior μ and the set $\Gamma(p)$ of the priors that reduce to p.[18] Notice that, by (22), the function ϕ plays a key role in determining the distance $I_t(\nu \parallel \mu)$ for all ν. In other words, here the role of ϕ is to induce a suitable distance of each prior ν relative to the reference prior μ.

The index $G(t, p)$ in (26) is, up to a shift t, exactly the distance $\min_{\nu\in\Gamma(p)} I_t(\nu \parallel \mu)$ between the reference prior μ and the set of priors $\Gamma(p)$.

By (26), the uncertainty averse version of the smooth representation is, for all $f \in \mathcal{F}$,

[18] In Probability Theory $\min_{\nu\in\Gamma(p)} I_t(\nu \parallel \mu)$ is called the I_t distance of μ from $\Gamma(p)$. An element of $\Gamma(p)$ where the minimum is achieved is called the projection of μ on $\Gamma(p)$ (see Csiszar [13]).

$$V(f) = \phi^{-1}\left(\int_{\Delta} \phi\left(\int_S u(f)\,dp\right) d\mu(p)\right) = \min_{p\in\Delta}\left(\int_S u(f)\,dp + \min_{\nu\in\Gamma(p)} I_{\int_S u(f)dp}(\nu \parallel \mu)\right). \tag{27}$$

This equality can be written in the following more interesting form.

Proposition 20 *For each $f \in \mathcal{F}$,*

$$V(f) = \min_{\nu\in\Delta^{\sigma}(\mathcal{B}(\Delta),\mu)} \left\{\int_{\Delta}\left(\int_S u(f)\,dq\right) d\nu(q) + I_{\int_{\Delta}\left(\int_S u(f)dq\right)d\nu(q)}(\nu \parallel \mu)\right\}. \tag{28}$$

This result gives the prior uncertainty interpretation of the smooth model. In fact, recall that without ambiguity the standard Bayesian model with a single prior μ is given by

$$\int_S u(f)\,dp = \int_{\Delta}\left(\int_S u(f)\,dq\right) d\mu(q). \tag{29}$$

The left hand side (lhs, for short) is the reduced form of the model, while the rhs is the "extensive" form. Under ambiguity, the decision maker has not enough information to quantify his beliefs with a single prior over the models in Δ. With this "prior uncertainty," he is only able to specify a reference prior μ on Δ that, for some reason, stands out among all possible priors. In the rhs of (28) all other possible priors ν in $\Delta^{\sigma}(\mathcal{B}(\Delta), \mu)$ are considered via their expected utilities

$$\int_S u(f)\,dp = \int_{\Delta}\left(\int_S u(f)\,dq\right) d\nu(q). \tag{30}$$

Each prior ν is then weighted through its distance

$$I_{\int_{\Delta}\left(\int_S u(f)dq\right)d\nu(q)}(\nu \parallel \mu) \tag{31}$$

from the reference prior μ. In particular, the smaller this distance is, the more relevant the prior ν is. The weight is minimal when $\nu = \mu$ and maximal when $I_{\int_{\Delta}\left(\int_S u(f)dq\right)d\nu(q)}(\nu \parallel \mu) = \infty$. In the latter case, the prior ν is "too distant" from the reference prior μ, and is therefore not taken into account. For example, when ϕ is the identity, then all priors ν distinct from μ are dismissed for this reason (see Example 14), and (28) reduces to the classic Bayesian equality (29).

Summing up, equality (28) thus shows that the smooth representation can be seen as a generalization of the classic Bayesian equality (29) that takes into account prior uncertainty. In fact, (28) considers all possible priors ν, evaluated through their expected utilities (30) and weighted through their distances (31) from the reference prior μ. Proposition 20 thus clarifies the prior uncertainty interpretation of the smooth representation, which is a novel insight that our analysis, and in particular Theorem 19, delivers.

5.2.3 Revealed Ambiguity

By showing what form Theorem 10 takes in the smooth case, the next important result confirms a key intuition of the smooth representation, that is, that the support $\operatorname{supp}\mu$ of μ is the set of probabilistic models that the decision maker considers relevant.

Theorem 21 *Let (u, ϕ, μ) be a smooth representation of $\succsim$. If $u(X) = \mathbb{R}$, ϕ is concave, and μ admits support,*[19] *then*

$$C^* = \operatorname{cl}(\operatorname{dom}_\Delta G) = \operatorname{cl}(\operatorname{co}(\operatorname{supp}\mu))$$

provided at least one of the following conditions holds:

(i) ϕ is bounded above;

(ii) $\lim_{t\to-\infty} \phi'_-(t) = \infty$.

Since a concave and strictly monotone function ϕ is unbounded from below, condition (i) suggests that, as t gets larger, utility losses carry more and more weight than utility gains. On the other hand, condition (ii) is automatically satisfied by order Orlicz functions. Inspection of the proof shows that the concavity of ϕ can be replaced by a weaker condition, still in the spirit of loss aversion. Hence, the equality $C^* = \operatorname{cl}(\operatorname{co}(\operatorname{supp}\mu))$ holds more generally even for some smooth representations that do not feature a concave ϕ.

5.2.4 Exponential Case and Overlap

Consider the important exponential case $\phi(t) = -e^{-\theta t}$, which corresponds to constant ambiguity aversion (see [34, p. 1866]). In this case we have the following version of Theorem 19, where $I_t(\cdot\|\mu)$ reduces to the relative entropy $R(\cdot\|\mu)$.

Corollary 22 *Let $u : X \to \mathbb{R}$ be an affine function with $u(X) = \mathbb{R}$, $\theta > 0$ a real number, and μ a countably additive Borel probability measure on Δ. The following conditions are equivalent:*

(i) $\left(u, -e^{-\theta(\cdot)}, \mu\right)$ is a smooth representation of $\succsim$,

(ii) (u, G) is an uncertainty averse representation of $\succsim$, where

$$G(t,p) = t + \frac{1}{\theta} \min_{\nu\in\Gamma(p)} R(\nu \parallel \mu) \qquad \forall (t,p) \in \mathbb{R}\times\Delta.$$

[19] See, e.g., [2, Ch. 12].

(iii) (u,c) *is a variational representation of* $\succsim$*, where* $c(p) = \frac{1}{\theta}\min_{\nu\in\Gamma(p)} R(\nu \parallel \mu)$ *for all* $p \in \Delta$*.*

Hence, here (27) and (28) become:

$$V(f) = -\frac{1}{\theta}\log \int_\Delta e^{-\theta \int_S u(f(s))dp(s)} d\mu(p) = \min_{p\in\Delta}\left\{\int_S u(f(s))\,dp(s) + \frac{1}{\theta}\min_{\nu\in\Gamma(p)} R(\nu \parallel \mu)\right\}$$
$$= \min_{\nu\in\Delta^\sigma(\mathcal{B}(\Delta),\mu)}\left\{\int_\Delta \left(\int_S u(f)\,dq\right) d\nu(q) + \frac{1}{\theta}R(\nu \parallel \mu)\right\}.$$

Corollary 22 thus shows what we already observed in the Introduction: the exponential case is thus both a smooth and a variational representation. Next we show that the exponential case is also, basically, the extent to which these two representations overlap.

Theorem 23 *Let* $u : X \to \mathbb{R}$ *be affine with* $u(X) = \mathbb{R}$ *and* $\phi : \mathbb{R} \to \mathbb{R}$ *be a strictly increasing and concave function. The triplet* (u, ϕ, μ) *represents a variational preference for all countably additive Borel probability measures* μ *on* Δ *if and only if* ϕ *is CARA.*

5.2.5 Quasi-Arithmetic Representation and Multiplier Preferences

We close this subsection by briefly considering preferences $\succsim$ that correspond to an objective function

$$V(f) = \phi^{-1}\left(\int_S (\phi \circ u)(f)\,dq\right) \qquad \forall f \in \mathcal{F} \tag{32}$$

where $u : X \to \mathbb{R}$ is an affine function, $\phi : \mathbb{R} \to \mathbb{R}$ is a strictly increasing and continuous function, and $q \in \Delta^\sigma$ is a (countably additive) probability on S. We call (u, ϕ, q) a *quasi-arithmetic representation* of $\succsim$ and we refer the interested reader to Strzalecki [54] for a recent discussion of this setting (see also [8] for some further properties of this setting).

When ϕ is the negative exponential $-e^{-\theta t}$, the representation (32) takes the variational form

$$V(f) = \min_{p\in\Delta^\sigma(q)}\left\{\int_S u(f)\,dp + \frac{1}{\theta}R(p \parallel q)\right\} \tag{33}$$

with the relative entropy $\theta^{-1}R(p \parallel q)$ as cost function. This variational representation corresponds to the Hansen and Sargent multiplier preferences ([32] and [31]). In particular, Strzalecki [54] provided behavioral conditions on variational preferences that characterize (33).

When ϕ is a general concave function, not necessarily exponential, the quasi-arithmetic representation (32) is uncertainty averse but, in general, no longer variational. The next result, based on the techniques that we just developed to represent smooth preferences, establishes the general uncertainty averse representation of (32), thus generalizing its variational representation (33) obtained for the ϕ exponential case.[20]

Theorem 24 *Let $u : X \to \mathbb{R}$ be an affine function with $u(X) = \mathbb{R}$, $\phi : \mathbb{R} \to \mathbb{R}$ a strictly increasing and concave function, and $q \in \Delta^\sigma$ a probability measure on S. The following conditions are equivalent:*

(i) (u, ϕ, q) is a quasi-arithmetic representation of $\succsim$,

(ii) (u, G) is an uncertainty averse representation of $\succsim$, where for each $t \in \mathbb{R}$,

$$G(t,p) = \begin{cases} t + I_t(p \parallel q) & \text{if } p \in \Delta^\sigma(q) \\ \infty & \text{else.} \end{cases} \tag{34}$$

In particular, $\phi(t) \approx -e^{-\theta t}$, with $\theta > 0$, if and only if $I_t(p \parallel q) = \theta^{-1} R(p \parallel q)$.

By (34), we thus have, for all $f \in \mathcal{F}$,

$$\phi^{-1}\left(\int_S (\phi \circ u)(f)\, dq\right) = \min_{p \in \Delta^\sigma(q)} \left\{\int u(f)\, dp + I_{\int u(f) dp}(p \parallel q)\right\}$$

which is the counterpart here of (28). In this case, $C^* = \mathrm{cl}\left(\bigcup_{t \in \mathbb{R}} \mathrm{dom} I_t(\cdot \parallel q)\right)$.

We close with a result, first proved by Strzalecki [54], that parallels Theorem 23.

Proposition 25 *Let $u : X \to \mathbb{R}$ be affine with $u(X) = \mathbb{R}$ and $\phi : \mathbb{R} \to \mathbb{R}$ be a strictly increasing and concave function. A triplet (u, ϕ, q) represents a variational preference for all probabilities $q \in \Delta^\sigma$ if and only if ϕ is CARA.*

In other words, the multiplier representation (33) is basically the overlap between variational and quasi-arithmetic representations.

5.3 Homothetic Preferences

Proposition 12 showed that variational preferences correspond to additively separable uncertainty indices. Next we study the multiplicatively separable case. A related model has been studied by Chateauneuf and Faro [10], as we detail below.

Behaviorally, this case turns out to be characterized by the following weak independence axiom with respect to a reference outcome x_* (think for example of the agent endowment).

[20] We omit the proof of this result because it is essentially an elementary version of the more complicated Theorem 19 and Corollary 22. Similarly, we omit the proof of Proposition 25, which is a simpler version of that of Theorem 23.

Axiom A. 10 (Homotheticity) *If $f, g \in \mathcal{F}$ and $\alpha, \beta \in (0,1]$,*

$$\alpha f + (1-\alpha)\, x_* \succsim \alpha g + (1-\alpha)\, x_* \Longrightarrow \beta f + (1-\beta)\, x_* \succsim \beta g + (1-\beta)\, x_*.$$

Relative to Axiom A.9, here the weights α and β can differ, while the constant act x_* is fixed. Axioms A.9 and A.10 can thus be regarded as symmetric weakenings of the Certainty Independence Axiom of Gilboa and Schmeidler [29] (see the discussion in [38, pp. 1454-1455]). In particular, a preference satisfies the Certainty Independence Axiom if and only if satisfies both Axioms A.9 and A.10.

Theorem 26 *Let $\succsim$ be an uncertainty averse preference that satisfies Axioms A.4-A.7 and (u, G) be an uncertainty averse representation of $\succsim$ such that $u(x_*) = 0$. The following conditions are equivalent:*

(i) $\succsim$ satisfies Axiom A.10;

(ii) there exist a nonempty, weak closed, and convex subset C of Δ and two functions $c_1, c_2 : C \to [0, \infty]$, such that*

(a) c_1 is concave and upper semicontinuous, with $0 < \inf_{p \in C} c_1(p) \leq \max_{p \in C} c_1(p) = 1$;

(b) c_2 is convex and lower semicontinuous, with $\min_{p \in C} c_2(p) = 1$;

(c) for all $(t, p) \in \mathbb{R} \times \Delta$,

$$G(t, p) = \begin{cases} \frac{t}{c_1(p)} & \text{if } t \geq 0 \text{ and } p \in C \\ \frac{t}{c_2(p)} & \text{if } t < 0 \text{ and } p \in C \\ \infty & \text{if } p \in \Delta \setminus C; \end{cases} \tag{35}$$

(iii) there exist $\gamma : \mathbb{R} \to \mathbb{R}_+$, with $\gamma(t) = 0$, if and only if $t = 0$, and $d_1, d_2 : \Delta \to (-\infty, \infty]$ such that, for all $(t, p) \in \mathbb{R} \times \Delta$,

$$G(t, p) = \begin{cases} \gamma(t)\, d_1(p) & \text{if } t \geq 0 \text{ and } p \in \Delta \\ \gamma(t)\, d_2(p) & \text{if } t < 0 \text{ and } p \in \Delta \end{cases}$$

with the convention $0 \cdot \infty = \infty$.

In this case, by (35), $C^* = C$. Moreover, Theorem 26 implies the following representation result.

Corollary 27 *Let $\succsim$ be a binary relation on $\mathcal{F}$. The following conditions are equivalent:*

(i) $\succsim$ *is uncertainty averse and satisfies Axioms A.4-A.7, and A.10;*

(ii) there exist an affine $u : X \to \mathbb{R}$*, with* $u(X) = \mathbb{R}$ *and* $u(x_*) = 0$*, a nonempty, weak* closed, and convex subset* C *of* Δ*, and two functions* $c_1, c_2 : C \to [0, \infty]$ *as in points (a) and (b) of Theorem 26 such that, for all* f *and* g *in* $\mathcal{F}$*,* $f \succsim g$ *if and only if*

$$\min_{p\in C}\left(\frac{\left(\int u(f)\,dp\right)^+}{c_1(p)} - \frac{\left(\int u(f)\,dp\right)^-}{c_2(p)}\right) \geq \min_{p\in C}\left(\frac{\left(\int u(g)\,dp\right)^+}{c_1(p)} - \frac{\left(\int u(g)\,dp\right)^-}{c_2(p)}\right). \tag{36}$$

In this case, u *is unique up to multiplication by a positive scalar,* C*,* c_1*, and* c_2 *are unique.*

For example, if $f(s), g(s) \succsim x_*$ for all $s \in S$, then (36) becomes:

$$f \succsim g \Longleftrightarrow \min_{p\in C}\frac{\int u(f)\,dp}{c_1(p)} \geq \min_{p\in C}\frac{\int u(g)\,dp}{c_1(p)}.$$

This is the specification studied by Chateauneuf and Faro [10], who assume the existence of a worst outcome with respect to which A.10 holds.

We close with two remarks. First, we already observed that a preference satisfies the Certainty Independence Axiom of Gilboa and Schmeidler [29] if and only if satisfies both Axioms A.9 and A.10. This means that a preference is both variational and homothetic if and only if is multiple priors. This can be seen also from the properties of the uncertainty aversion indices. In fact, by (20) and (35), an index G is both variational and homothetic if:

$$\begin{array}{ll} t + c(p) = \frac{t}{c_1(p)} & \text{if } t \geq 0 \text{ and } p \in C \\ t + c(p) = \frac{t}{c_2(p)} & \text{if } t < 0 \text{ and } p \in C \\ t + c(p) = \infty & \text{if } p \in \Delta \setminus C. \end{array}$$

It is easy to check that the unique solution is $c(p) = 0$ and $c_1(p) = c_2(p) = 1$ for all $p \in C$, and $c(p) = \infty$ if $p \notin C$. We thus get

$$f \succsim g \Longleftrightarrow \min_{p\in C}\int u(f)\,dp \geq \min_{p\in C}\int u(g)\,dp \tag{37}$$

which is the multiple priors criterion. Notice that for fixed u and C, by Proposition 6, the agent using criterion (37) is the most uncertainty averse of those using criterion (36).

Second, observe that in the proof of Corollary 27 we show that, when Σ is a σ-algebra, then $\succsim$ satisfies Axiom A.8 if and only if there is $q \in \Delta^\sigma$ such that $C \subseteq \Delta^\sigma(q)$.

6 Outline of the Main Proof

In this section we outline the arguments used in the proofs of Theorems 3 and 5, our main representation results. Assume for convenience that S is finite and take $\Sigma = 2^S$. This makes our setting finite dimensional since the collection of acts $\mathcal{F}$ can be identified with X^S and the utility profile $u(f)$ of each act corresponds to the vector $(u(f(s)))_{s\in S}$ of $\mathbb{R}^S$. We will also assume the unboundedness Axiom A.6, which guarantees $u(X) = \mathbb{R}$.

By the decision theoretic arguments of Lemmas 60, 61, and 63 in Appendix A (whose proofs are unaffected by the dimensionality of the setting), a preference $\succsim$ on X^S satisfies Axioms A.1-A.6 if and only if there exists an affine onto $u : X \to \mathbb{R}$ and a function $I : \mathbb{R}^S \to \mathbb{R}$ normalized, monotone, quasiconcave, and continuous such that

$$f \succsim g \Longleftrightarrow I(u(f)) \geq I(u(g)). \tag{38}$$

Moreover, u is cardinally unique, and, given u, there is a unique normalized I that satisfies (38).

Now, for any t in $\mathbb{R}$ and p in the probability simplex Δ, let $B(t,p) = \left\{v \in \mathbb{R}^S : p \cdot v \leq t\right\}$ and $G(t,p) = \sup_{v\in B(t,p)} I(v)$. The *formal* analogy with budget sets and indirect utilities of consumer theory is evident. Clearly, $t \leq t'$ implies $B(t,p) \subseteq B(t',p)$ for all p, and $B(\lambda(t,p) + (1-\lambda)(t',p')) \subseteq B(t,p) \cup B(t',p')$ for all t, t', p, p' and all $\lambda \in [0,1]$. Therefore $G(t,p)$ is increasing in the first variable and jointly quasiconvex.

For each v_0 in $\mathbb{R}^S$ and all p in Δ, $v_0 \in B(p \cdot v_0, p)$ which implies $I(v_0) \leq G(p \cdot v_0, p)$ and

$$I(v_0) \leq \inf_{p\in\Delta} G(p \cdot v_0, p).$$

On the other hand, continuity, quasiconcavity, and monotonicity of I imply that $C = \{I > I(v_0)\}$ is an open convex set such that $v_0 \notin C = C + \mathbb{R}^S_+$. A Separating Hyperplane Theorem guarantees that there exists $p_0 \in \Delta$ such that $p_0 \cdot v_0 < p_0 \cdot v$ for all $v \in C$. In particular, $B(p_0 \cdot v_0, p_0) \subseteq C^c$, that is, $G(p_0 \cdot v_0, p_0) \leq I(v_0)$. Hence,

$$I(v_0) = \min_{p\in\Delta} G(p \cdot v_0, p).$$

This, together with (38) delivers formula (9) and the cardinal uniqueness of u. To complete the proof that $G \in \mathcal{G}(\mathbb{R} \times \Delta)$ is enough to observe that I is normalized (that is, decision theoretically, that the decision maker is a von Neumann-Morgenstern expected utility maximizer when only risk is involved).

Lower semicontinuity of G is proved in Lemma 32 of Appendix A. As to uniqueness, we just sketch the main points of the proof. Let H be another lower semicontinuous element of $\mathcal{G}(\mathbb{R} \times \Delta)$ such that

$$I(v) = \min_{p\in\Delta} H(p \cdot v, p) \tag{39}$$

indeed, we should consider an element H such that (9) holds, but it is possible to reduce this case to (39). Fix $p_0 \in \Delta$ and $t_0 \in \mathbb{R}$, then

$$G(t_0, p_0) = \sup_{v \in B(t_0,p_0)} I(v) = \sup_{v \in B(t_0,p_0)} \inf_{p \in \Delta} H(p \cdot v, p)$$

a nontrivial (see [7] for details) minimax argument allows to exchange the sup and the inf delivering

$$G(t_0, p_0) = \inf_{p \in \Delta} \sup_{v \in B(t_0,p_0)} H(p \cdot v, p). \tag{40}$$

It remains to study the program

$$\sup_{v \in B(t_0,p_0)} H(p \cdot v, p).$$

If $p \neq p_0$, they generate different hyperplanes through 0, that is, there exists $v_0 \in \mathbb{R}^S$ such that $p_0 \cdot v_0 = 0$ and $p \cdot v_0 \neq 0$. Call e the unit vector of $\mathbb{R}^S$, $p_0 \cdot t_0 e = t_0$ and the straight line $\{\alpha v_0 + t_0 e : \alpha \in \mathbb{R}\}$ is included in $B(t_0, p_0)$. Thus

$$\sup_{v \in B(t_0,p_0)} H(p \cdot v, p) \geq \sup_{\alpha \in \mathbb{R}} H(\alpha p \cdot v_0 + t_0, p) = \sup_{t \in \mathbb{R}} H(t, p) = \infty$$

where the last equality follows from normalization of the elements of $\mathcal{G}(\mathbb{R} \times \Delta)$.

Else, if $p = p_0$, since H is increasing in the first component, we have

$$\sup_{v \in B(t_0,p_0)} H(p_0 \cdot v, p_0) = H(t_0, p_0).$$

In conclusion, $\sup_{v \in B(t_0,p_0)} H(p \cdot v, p) = H(t_0, p_0)$ if $p = p_0$ and ∞ otherwise. Hence,

$$\inf_{p \in \Delta} \sup_{v \in B(t_0,p_0)} H(p \cdot v, p) = H(t_0, p_0)$$

which together with (40) delivers $G = H$.

When $u(X) \subseteq \mathbb{R}$, only the usual minimax inequality holds. Thus, equality in (40) is replaced by minorization, and only the minimality of G follows.

It should be remarked that the main simplifying assumption here is *not* the finite dimensionality of the setting (which only eliminates some functional analytic subtleties), but rather the assumption $u(X) = \mathbb{R}$, which also in the finite dimensional case is hard to weaken. To overcome this problem some extension techniques are used (see Theorem 37). Moreover, the passage from the *non topological* assumptions of the preference axioms to the strong continuity properties of I is also nontrivial and led us to study thoroughly the continuity properties of monotone quasiconcave functions (see Section A.3 of Appendix A). A complete duality for this class of functions is studied in [7] and is very different from the classical Fenchel duality on which the results of [38] rest. This duality already proved its usefulness in other applications such as risk management (see Cerreia-Vioglio, Maccheroni, Marinacci, and Montrucchio [9]).

7 A Game Theoretic Perspective

As mentioned in the Introduction, our setting admits a game against Nature interpretation, where decision makers view themselves as playing a zero-sum game against (a malevolent) Nature. Here f and p become, respectively, the strategies of the decision maker and of Nature, and the interpretation of the axioms and of the results has to be suitably modified.

Begin with the representation: by Theorem 3 for every uncertainty averse preference there exists a (suitably unique) game against Nature

$$\Gamma(f,p) = G\left(\int u(f)\,dp, p\right) \qquad \forall (f,p) \in \mathcal{F} \times \Delta$$

such that the agent behaves as if he were playing this game against Nature, and conversely for every such game there is a unique corresponding uncertainty averse preference.

The structure of the game reflects the existence of two sources of uncertainty in the Anscombe-Aumann model (see also Strzalecki [54]):

- $\int u(f)\,dp$ captures the objective risk preferences of the agent (once Nature has chosen p, the expected utility of f is what matters to the agent); notice that, for fixed p, $G(\cdot, p)$ is an increasing transformation of the expected utility.
- G captures the presence of subjective uncertainty and has the standard properties of convexity and continuity of a zero-sum game.
- The minimum – that is, the adoption of a maxmin strategy – captures the aversion of the decision maker to such uncertainty.

In this perspective, the reason why in Axiom A.3 the decision maker prefers to randomize among indifferent acts is because this makes more difficult for Nature (which has no control on the randomizing device) to reply.

In turn, the fact that Nature has no control on the randomizing device is captured in the representation by the normalization of G, that is, by the fact that, for any constant act x,

$$V(x) = \inf_{p\in\Delta} G\left(\int u(x)\,dp, p\right) = \inf_{p\in\Delta} G(u(x), p) = u(x).$$

This normalization is natural: given a decision maker's act f, Nature can affect the relative likelihood of the act's outcomes by choosing a probabilistic model p, unless f is constant, in which case Nature has no power.

In a similar vein, if the decision maker $\succsim_2$ prefers an uncertain act f over a constant one x whenever also $\succsim_1$ does, here this means that $\succsim_2$ is less worried than $\succsim_1$ about Nature's ability to impair his acts' outcomes. Accordingly, Proposition 6 on comparative uncertainty aversion here says that $\succsim_1$ is more uncertainty averse than $\succsim_2$ if and only if $u_1 \approx u_2$ and $\Gamma_1 \leq \Gamma_2$ (provided $u_1 = u_2$). In other words, relative to $\succsim_2$, the decision maker $\succsim_1$ behaves as if he is believing to face a more powerful Nature, that is, a Nature that incurs in smaller losses for her actions.

By Theorem 8, in this game theoretic perspective probabilistic sophistication corresponds to a situation where for every move f of the agent the loss incurred by Nature in choosing p is only determined by how much the masses $dp(s)$ are scattered w.r.t the masses $dq(s)$, that is, by how much the reference probability is "warped." For example, if q is uniform, Nature losses will be independent of permutations of the states. In sum, the decision maker is probabilistically sophisticated if and only if he thinks that also Nature is.

Finally, by Theorem 10, the set C^* can be viewed as the closure of Nature's conceivable actions (for, she will never play an action inducing an infinite loss). In particular,

$$f \succsim^* g \Longleftrightarrow \int u(f)\, dp \geq \int u(g)\, dp \qquad \forall p \in \mathrm{dom}_\Delta G$$

means that if f is unambiguously preferred to g, then, when regarded as strategies, f weakly dominates g in the game. Similarly, Proposition 11 implies that $\pi_u(f)$ is the set of best replies for Nature to act f.

A Quasiconcave Monotone Functionals

In this Appendix we report the properties of a duality notion for monotone quasiconcave functionals on which the results of the paper rest. This topic is studied in detail in [7], to which we refer the interested reader.[21]

Notation 28 *In this section and in the next one we denote by X (resp., $g : X \to [-\infty, \infty]$) an ordered vector space (resp., an extended valued function).*

This makes it easier to refer to [7]. See [37] and [2, Ch. 9] for all notions on ordered vector spaces used here.

A.1 Preliminaries

A.1.1 Set Up

The Space and its Geometry

[21] The nontrivial proofs that we omit here can be found in [7].

Assumption 1 $(X, \|\cdot\|, \geq)$ *is a normed Riesz space with order unit* e *and* $\|\cdot\|$ *is its supnorm, i.e.* $\|x\| = \inf\{\alpha \in \mathbb{R} : |x| \leq \alpha e\}$ *for all* $x \in X$.

Recall that any norm on a normed Riesz space with order unit is equivalent to the supnorm induced by the unit.

The most relevant example for this paper is the function space $B_0(\Sigma)$, with order unit 1_S. $B_0(\Sigma)$ also has the following important property: for every ideal J of $B_0(\Sigma)$, the quotient space $B_0(\Sigma)/J$ is Archimedean. Normed Riesz spaces with this property are called *hyper-Archimedean*, and every hyper-Archimedean space is actually Riesz isomorphic to suitable space $B_0(\Sigma)$ (see, e.g., [37, Thm. 37.7]).

If $y, z \in X$, $[y, z]$ is the order interval $\{x \in X : y \leq x \leq z\}$. Notice that the closed unit ball of X coincides with

$$[-e, e] = \{x \in X : -e \leq x \leq e\}. \tag{41}$$

Denoting by X_+ and X_- the positive and negative cones in X, then the positive and negative unit balls are $[-e, e] \cap X_+ = [0, e]$ and $[-e, e] \cap X_- = [-e, 0]$. A subset Y of X is *lower open* (resp., *upper open*) if for all $y \in Y$ there exists $\varepsilon > 0$ such that $[y - \varepsilon e, y] \subseteq Y$ (resp., $[y, y + \varepsilon e] \subseteq Y$). Clearly, open sets are lower and upper open (but, there are subsets of $\mathbb{R}^2$ which are lower and upper open, without being open).

For every $x \in X$, set

$$\operatorname{ess\,sup}(x) = \inf\{\alpha \in \mathbb{R} : x \leq \alpha e\} \quad \text{and} \quad \operatorname{ess\,inf}(x) = -\operatorname{ess\,sup}(-x).$$

By definition of supnorm, $\|\cdot\| = \operatorname{ess\,sup}(|\cdot|)$. For any interval T of the real line, set

$$X(T) = \{x \in X : [\operatorname{ess\,inf}(x), \operatorname{ess\,sup}(x)] \subseteq T\}.$$

It is easy to check that $X(T)$ is convex, and either lower open (if and only if $\inf T \notin T$) or upper open (if and only if $\sup T \notin T$) or it is an order interval ($[(\inf T)e, (\sup T)e]$). Moreover, it is open if and only if T is open. If $X = B_0(\Sigma)$, then $X(T) = B_0(\Sigma, T)$ is the set of functions in $B_0(\Sigma)$ whose range is contained in T.

We denote by X^* the topological dual of X. Elements of X^* are usually denoted by ξ, and $\langle \xi, x\rangle$, with $x \in X$, denotes the duality pairing $\xi(x)$. X_+^* the set of all positive functionals in X^*. Notice that, by (41), $\|\xi\| = \langle \xi, e\rangle$ for all $\xi \in X_+^*$. In particular the set

$$\Delta = \left\{\xi \in X_+^* : \|\xi\| = 1\right\}$$

is (and weak* compact and) convex since it coincides with $\left\{\xi \in X_+^* : \langle \xi, e\rangle = 1\right\}$.

Assumption 2 Δ *is equipped with the weak* topology.*

A subset C of X is *evenly convex* if it is the intersection of a family of open half spaces.[22] Evenly convex sets are convex, and intersections of evenly convex sets are evenly convex. By standard separation results, both open convex sets and closed convex sets are then evenly convex. Moreover, for every interval T of $\mathbb{R}$, $X(T)$ is evenly convex.

Functions

If $Y \subseteq X$, $g : Y \to [-\infty, \infty]$, and $a \in [-\infty, \infty]$, we set $\{g \geq a\} = \{y \in Y : g(y) \geq a\}$, the sets $\{g > a\}$, $\{g \leq a\}$, and $\{g < a\}$ are defined in the same way.

For functions $g : X \to [-\infty, \infty]$, the relevant notion of effective domain, $\mathrm{dom}(g)$ depends on whether we consider the hypograph or the epigraph of g. In the former case we have $\mathrm{dom}(g) = \{g > -\infty\}$, while in the latter case we have $\mathrm{dom}(g) = \{g < \infty\}$. For functions $g : X \to [-\infty, \infty)$ it is natural to consider hypographs, and so $\mathrm{dom}(g) = \{g > -\infty\}$. Symmetrically, we have $\mathrm{dom}(g) = \{g < \infty\}$ for functions $g : X \to (-\infty, \infty]$. In any other case the definition of $\mathrm{dom}(g)$ will be explicitly given.

A function $g : X \to [-\infty, \infty]$ is:

- *monotone* if $x \geq y$ implies $g(x) \geq g(y)$;
- *evenly quasiconcave* if the sets $\{g \geq \alpha\}$ are evenly convex for all $\alpha \in \mathbb{R}$;
- *evenly quasiconvex* if the sets $\{g \leq \alpha\}$ are evenly convex for all $\alpha \in \mathbb{R}$;
- *positively homogeneous* if $g(\lambda x) = \lambda g(x)$ for all $\lambda > 0$ and $x \in X$;
- *normalized* if $g(\alpha e) = \alpha$ for all $\alpha \in \mathbb{R}$;
- *translation invariant* if $g(x + \alpha e) = g(x) + \alpha$ for all $\alpha \in \mathbb{R}$.

Clearly, evenly quasiconcave functions are quasiconcave. Moreover, both lower and upper semicontinuous quasiconcave functions on X are evenly quasiconcave.

Observe that when g is positively homogeneous on X, then $g(0) = \lambda g(0)$ for all $\lambda > 0$, so that either $g(0) = \pm\infty$ or $g(0) = 0$. In particular, $g(0) = 0$ if it is finite.

If g is defined on a subset Y of X the above definitions remain unchanged with the additional requirement that all the arguments of $g(\cdot)$ belong to Y.[23]

If $\{x_n\}$ is a sequence in X, write $x_n \nearrow x$ (resp., $x_n \searrow x$) if it is increasing (resp., decreasing) and it converges to x in norm. A function $g : Y \to \mathbb{R}$ is:

[22] With the convention that such intersection is X if the family is empty. The notion of even convexity and its basic properties are due to Fenchel [22].

[23] For example, positive homogeneity becomes: $g(\lambda x) = \lambda g(x)$ for all $\lambda > 0$ and $x \in X$ such that $\lambda x, x \in Y$.

- *left (sequentially) continuous* at $x \in Y$ if $\{x_n\} \subseteq Y$ and $x_n \nearrow x$ implies $g(x_n) \to g(x)$;
- *right (sequentially) continuous* at $x \in Y$ if $\{x_n\}_n \subseteq Y$ and $x_n \searrow x$ implies $g(x_n) \to g(x)$.

Upper (and Lower) Semicontinuous Envelopes

Given $x \in X$, denote by $\mathcal{N}_x$ the set of all neighborhoods of x in X. Given a function $g : X \to [-\infty, \infty]$, its *upper semicontinuous envelope* $g^+ : X \to [-\infty, \infty]$ is defined by $g^+(x) = \inf_{U \in \mathcal{N}_x} \sup_{y \in U} g(y)$ for all $x \in X$ (see [14, Ch. 3]); and hence

$$\{g^+ \geq \alpha\} = \bigcap_{\beta < \alpha} \overline{\{g > \beta\}} \qquad \forall \alpha \in \mathbb{R}. \tag{42}$$

Moreover, g^+ is the least upper semicontinuous function on X that pointwise dominates g.

Lemma 29 *If $g : X \to [-\infty, \infty]$ is monotone, then g^+ is monotone and $g^+(x) = \inf_n g(x_n)$ for all $x \in X$ and every sequence x_n such that $x_n \to x$ and $x_n > x$ for all $n \in \mathbb{N}$.*[24] *Moreover, g^+ is quasiconcave provided g is.*

Proof. Let $x \in X$. For each $n \geq 1$, set $V_n = [2x - x_n, x_n] = [x - e_n, x + e_n]$, where $e_n = x_n - x$ for all $n \in \mathbb{N}$. Belonging to the interior of X_+, e_n is an order unit for all $n \in \mathbb{N}$, and $e_n \to 0$. In particular, $V_n \in \mathcal{N}_x$ for all $n \in \mathbb{N}$. Therefore, $\inf_{U \in \mathcal{N}_x} \sup_{y \in U} g(y) \leq \inf_n \sup_{y \in V_n} g(y)$. Moreover, since $e_n \to 0$, for each $U \in \mathcal{N}_x$ there is $n_U \in \mathbb{N}$ such that $V_{n_U} \subseteq U$,[25] and we also have $\sup_{y \in U} g(y) \geq \sup_{y \in V_{n_U}} g(y) \geq \inf_n \sup_{y \in V_n} g(y)$. Then $\inf_{U \in \mathcal{N}_x} \sup_{y \in U} g(y) \geq \inf_n \sup_{y \in V_n} g(y)$, and $g^+(x) = \inf_n \sup_{y \in V_n} g(y)$. By monotonicity of g, $\sup_{y \in V_n} g(y) = g(x_n)$ and $g^+(x) = \inf_n g(x_n)$.

If $z \in X$ and $x \leq z$, then $g(x + n^{-1}e) \leq g(z + n^{-1}e)$ for all $n \in \mathbb{N}$, whence $g^+(x) = \inf_n g(x + n^{-1}e) \leq \inf_n g(z + n^{-1}e) = g^+(z)$, thus g^+ is monotone.

Finally, if g is quasiconcave, (42) implies that g^+ as well is quasiconcave. ■

Totally analogous results hold for lower semicontinuity: Given a function $g : X \to [-\infty, \infty]$, its *lower semicontinuous envelope* $g^- : X \to [-\infty, \infty]$ is defined by $g^-(x) = \sup_{U \in \mathcal{N}_x} \inf_{y \in U} g(y)$ for all $x \in X$; and hence $\{g^- > \alpha\} = \bigcup_{\beta > \alpha} \{g > \beta\}^\circ$ for all $\alpha \in \mathbb{R}$. Moreover, g^- is the greatest lower semicontinuous function on X that is pointwise dominated by g.

[24] $x_n > x$ means that $x_n - x$ belongs to the interior of X_+ (while $x_n \gneq x$ means that $x_n \geq x$ and $x_n \neq x$).

[25] There exists $\delta > 0$ such that $[x - \delta e, x + \delta e] \subseteq U$, but $e_n \to 0$ implies that eventually $-e_n, e_n \subseteq [-\delta e, \delta e]$, and $[x - e_n, x + e_n] \subseteq [x - \delta e, x + \delta e] \subseteq U$.

Lemma 30 *If $g : X \to [-\infty, \infty]$ is monotone, then g^- is monotone and $g^-(x) = \sup_n g(x_n)$ for all $x \in X$ and every sequence x_n such that $x_n \to x$ and $x > x_n$ for all $n \in \mathbb{N}$. Moreover, g^- is quasiconcave provided g is.*

A.1.2 A Key Auxiliary Function

Let $\emptyset \neq Y \subseteq X$ and $g : Y \to [-\infty, \infty]$. Set

$$G_\xi(t) = \sup\{g(x) : x \in Y \text{ and } \langle \xi, x \rangle \leq t\}$$

for all $(t, \xi) \in \mathbb{R} \times \Delta$, with the usual convention $\sup \emptyset = -\infty$.

This function, plays a key role in what follows, it can take values on $[-\infty, \infty]$. Since, for our analysis, the set where it can take on value ∞ is more relevant than that where it takes on value $-\infty$, throughout the appendix we set $\operatorname{dom}(G_\xi) = \{G_\xi < \infty\}$.

The function G_ξ is monotone, thus, denoting by G_ξ^+ the upper semicontinuous envelope of G_ξ, by Lemma 29, $G_\xi^+(t) = \inf\{G_\xi(t') : t' > t\}$. The next lemmas give some basic properties of the function G_ξ (the proofs, when omitted, can be found in [7]).

Lemma 31 *For any function $g : Y \to [-\infty, \infty]$, the map $(t, \xi) \mapsto G_\xi(t)$ is quasiconvex over $\mathbb{R} \times \Delta$. Moreover, $\lim_{t \to \infty} G_\xi(t) = \sup_{\zeta \in \Delta} \sup_{t \in \mathbb{R}} G_\zeta(t) = \sup_{x \in Y} g(x)$ for all $\xi \in \Delta$.*

Lemma 32 *Let Y be lower open and $g : Y \to [-\infty, \infty]$ be monotone and lower semicontinuous. Then, the map $(t, \xi) \mapsto G_\xi(t)$ is lower semicontinuous on $\mathbb{R} \times \Delta$.*

Proof. Let $\lambda \in \mathbb{R}$ and $(\bar{t}, \bar{\xi}) \in \mathbb{R} \times \Delta$ be such that $G_{\bar{\xi}}(\bar{t}) > \lambda$. We want to show that $G_\xi(t) > \lambda$ for all (t, ξ) in a suitable neighborhood of $(\bar{t}, \bar{\xi})$.

Since $\sup_{y \in Y : \langle \bar{\xi}, y \rangle \leq \bar{t}} g(y) > \lambda$, there is $y_0 \in Y$ such that $\langle \bar{\xi}, y_0 \rangle \leq \bar{t}$ and $g(y_0) > \lambda$. Since Y is lower open, eventually the sequence $y_n = y_0 - n^{-1} e$ belongs to Y and $y_n \nearrow y_0$. As g is lower semicontinuous, there exists $\bar{n} \in \mathbb{N}$ such that $y_{\bar{n}} \in Y$ and $g(y_{\bar{n}}) > \lambda$. Moreover, $\langle \bar{\xi}, y_{\bar{n}} \rangle = \langle \bar{\xi}, y_0 \rangle - \bar{n}^{-1} \langle \bar{\xi}, e \rangle \leq \bar{t} - \delta$ for $\delta = \bar{n}^{-1}$.

The set $U = \{\xi \in \Delta : \langle \xi, y_{\bar{n}} \rangle < \langle \bar{\xi}, y_{\bar{n}} \rangle + \delta/2\}$ is open in the induced weak* topology of Δ, and for all $(t, \xi) \in (\bar{t} - \delta/2, \infty) \times U$ we have $\langle \xi, y_{\bar{n}} \rangle \leq \langle \bar{\xi}, y_{\bar{n}} \rangle + \delta/2 \leq \bar{t} - \delta + \delta/2 = \bar{t} - \delta/2 < t$. Hence, $G_\xi(t) = \sup_{y \in Y : \langle \xi, y \rangle \leq t} g(y) \geq g(y_{\bar{n}}) > \lambda$, as wanted. ■

Remark 33 *In particular, for all $\xi \in \Delta$, the map $t \mapsto G_\xi(t)$ is lower semicontinuous and monotone, therefore it is left continuous.*

In the next Lemmas we assume $Y = X$.

Lemma 34 *If $g : X \to [-\infty, \infty]$ is monotone, then $G_\xi(t) = \sup\{g(x) : x \in X \text{ and } \langle \xi, x\rangle = t\}$ for all $(t, \xi) \in \mathbb{R}\times\Delta$.*

Lemma 35 *Let $h : X \to [-\infty, \infty]$, $\varphi : \overline{h(X)} \to [-\infty, \infty]$ be extended-valued continuous and monotone, and $g = \varphi \circ h$. Then, $G_\xi(t) = \varphi(H_\xi(t))$ for all $(t, \xi) \in \mathbb{R} \times \Delta$.*

A.2 General Representation

A.2.1 A Theorem of de Finetti and its Extension

The next result shows that a function g can be recovered from the scalar function $G_\xi(t)$ as long as g is quasiconcave. Here we only consider the monotone case, and we refer the reader to [7] for a general version and for a proof. An early version of this result can be found in de Finetti [16, p. 178], while a closely related general formulation can be found in [44, Theorem 2.6]. Notice that versions of this result play an important role in microeconomic duality theory (see, e.g., Diewert [18]).

Theorem 36 *A function $g : X \to [-\infty, +\infty]$ is evenly quasiconcave and monotone if and only if*

$$g(x) = \inf_{\xi\in\Delta} G_\xi(\langle \xi, x\rangle) \qquad \forall x \in X. \tag{43}$$

Moreover, if g is lower semicontinuous, then the infimum in (43) is attained, while if g is upper semicontinuous, then G_ξ can be replaced with G_ξ^+.

The next result considers the representation (43) for a monotone function defined on a subset Y.

Theorem 37 *Let $g : Y \to \mathbb{R}$ be a quasiconcave and monotone function defined on a convex subset Y of X. Then,*

$$g(y) = \inf_{\xi\in\Delta} G_\xi(\langle \xi, y\rangle) \qquad \forall y \in Y \tag{44}$$

provided at least one of the following conditions hold:

(i) g is lower semicontinuous and Y is lower open;

(ii) g is upper semicontinuous and Y is either upper open or it is an order interval.

Moreover, under condition (i) the infimum is attained.

Proof. (i) Suppose that g is lower semicontinuous and that Y is lower open. We want to prove (44) with min in place of inf. The function $\hat{g} : X \to [-\infty, \infty]$ defined by

$$\hat{g}(x) = \sup \{g(y) : Y \ni y \leq x\} \tag{45}$$

is the minimal monotone extension of g to X (with the usual convention $\sup \emptyset = -\infty$).

Assume first that Y is open. Since $\{x \in X : \hat{g}(x) > t\} = \{y \in Y : g(y) > t\} + X_+$ for all $t \in \mathbb{R}$, the function $\hat{g}$ is quasiconcave and lower semicontinuous. By Theorem 36, $\hat{g}(x) = \min_{\xi \in \Delta} \hat{G}_\xi(\langle \xi, x \rangle)$ for all $x \in X$. Hence, given $y \in Y$, there is $\xi_y \in \Delta$ such that $\hat{g}(y) = \hat{G}_{\xi_y}(\langle \xi_y, y \rangle) \geq G_{\xi_y}(\langle \xi_y, y \rangle) \geq g(y) = \hat{g}(y)$. Hence, $g(y) = \min_{\xi \in \Delta} G_\xi(\langle \xi, y \rangle)$.

If Y is only lower open, consider the lower semicontinuous envelope $\hat{g}^-$ of $\hat{g}$, simply denoted by $\tilde{g}$. Since $\hat{g}$ is monotone and quasiconcave so is $\tilde{g}$ (see Lemma 30). Moreover, $\tilde{g}$ extends g. In fact, for all $y \in Y$, $\tilde{g}(y) = \sup_n \hat{g}(y - n^{-1}e) = \lim_n \hat{g}(y - n^{-1}e) = \lim_n g(y - n^{-1}e)$ since eventually $y - n^{-1}e \in Y$, and, by monotonicity and lower semicontinuity of g on Y, $g(y) \geq \lim_n g(y - n^{-1}e) \geq g(y)$. By proceeding as in the first part of the proof, we can then prove that (44) holds.[26]

(ii) Suppose that g is upper semicontinuous and that Y is either upper open or it is an order interval $[w, z]$. Consider the function $\hat{g} : X \to [-\infty, \infty]$ defined in (45). From point (i) we know that $\hat{g}$ is the minimal monotone extension of g to X, and that $\hat{g}$ is quasiconcave. By Lemma 29, its upper semicontinuous envelope $\hat{g}^+$ is monotone and quasiconcave too. Denote it by $\bar{g}$. Next we show that $\bar{g}$ extends g.

- If Y is upper open. Let $y \in Y$, then $\bar{g}(y) = \inf_n \hat{g}(y + n^{-1}e) = \lim_n \hat{g}(y + n^{-1}e) = \lim_n g(y + n^{-1}e)$ since eventually $y + n^{-1}e \in Y$, and, by monotonicity and lower semicontinuity of g on Y, $g(y) \leq \lim_n g(y + n^{-1}e) \leq g(y)$.

- If $Y = [w, z]$, for some $w, z \in X$. We show that $\hat{g}$ is upper semicontinuous on X, then $\hat{g} = \hat{g}^+ = \bar{g}$, and $\bar{g}$ extends g, since $\hat{g}$ does. Let $x \in X_+ + w$. For all $y \in Y$ such that $y \leq x$, then $y \leq x \wedge z \leq x$ and $w \leq y \leq x \wedge z \leq z$ imply that $x \wedge z \in Y$ and $g(y) \leq g(x \wedge z)$, thus $g(y) \leq g(x \wedge z) \leq \hat{g}(x)$. Since this is true for all $y \in Y$ such that $y \leq x$, then $\hat{g}(x) = \sup\{g(y) : Y \ni y \leq x\} \leq g(x \wedge z) \leq \hat{g}(x)$; but the choice of x was arbitrary, hence $\hat{g}(x) = g(x \wedge z)$ for all $x \in X_+ + w$. If $x_n, x \in X_+ + w$ and $x_n \to x$, then $x_n \wedge z \to x \wedge z$ and $\limsup_n \hat{g}(x_n) = \limsup_n g(x_n \wedge z) \leq g(x \wedge z) = \hat{g}(x)$. This shows that $\hat{g}$ is upper semicontinuous on the closed set $X_+ + w$. Together with $\hat{g}(x) = -\infty$ for all $x \notin X_+ + w$, this shows that $\hat{g}$ is upper semicontinuous on X.

[26] By Theorem 36, $\tilde{g}(x) = \min_{\xi \in \Delta} \tilde{G}_\xi(\langle \xi, x \rangle)$, for all $x \in X$. Hence, given $y \in Y$, there is $\xi_y \in \Delta$ such that $\tilde{g}(y) = \tilde{G}_{\xi_y}(\langle \xi_y, y \rangle) \geq G_{\xi_y}(\langle \xi_y, y \rangle) \geq g(y) = \tilde{g}(y)$. Hence, $g(y) = \min_{\xi \in \Delta} G_\xi(\langle \xi, y \rangle)$.

For all $(t,\xi)\in\mathbb{R}\times\Delta$, $G_{\xi}(t)=\sup\{g(x):x\in Y,\langle\xi,x\rangle\leq t\}=\sup\{\bar{g}(x):x\in Y,\langle\xi,x\rangle\leq t\}$ $\leq\sup\{\bar{g}(x):x\in X,\langle\xi,x\rangle\leq t\}=\bar{G}_{\xi}(t)$. By Theorem 36, for all $y\in Y$, $g(y)=\bar{g}(y)$ $=\inf_{\xi\in\Delta}\bar{G}_{\xi}(\langle\xi,y\rangle)\geq\inf_{\xi\in\Delta}G_{\xi}(\langle\xi,y\rangle)\geq g(y)$, as desired. ■

Corollary 38 *Let $g:X(T)\to\mathbb{R}$ be quasiconcave and monotone. If g is continuous, then $g(x)=\inf_{\xi\in\Delta}G_{\xi}(\langle\xi,x\rangle)$ for all $x\in X(T)$. In particular, the infimum are attained if T is lower open.*

A.2.2 Concavity

The next two corollaries of Theorem 36, proved in [7], give some characterizations of concavity. Here $g^{*}:X^{*}\to[-\infty,\infty]$ denotes the classic (concave) Fenchel conjugate of g, given by $g^{*}(\xi)=\inf_{x\in X}\{\langle\xi,x\rangle-g(x)\}$ for all $\xi\in X^{*}$.

Corollary 39 *Let $g:X\to\mathbb{R}$ be evenly quasiconcave and monotone. The following facts are equivalent:*

(i) g is concave;

(ii) G_{ξ} is concave for each $\xi\in\Delta$;

(iii) $G_{\xi}(t)=\inf_{\lambda\in\mathbb{R}_{+}}\{\lambda t-g^{*}(\lambda\xi)\}$ for each $(t,\xi)\in\mathbb{R}\times\Delta$.

In particular, $\mathrm{dom}(G_{\xi})\in\{\emptyset,\mathbb{R}\}$ *for all $\xi\in\Delta$.*

Next we consider normalized functions.

Corollary 40 *Let $g:X\to[-\infty,\infty]$ be monotone and evenly quasiconcave. g is normalized if and only if $\inf_{\xi\in\Delta}G_{\xi}(t)=t$ for all $t\in\mathbb{R}$. Moreover, the following properties are equivalent:*

(i) g is concave and normalized;

(ii) g is translation invariant and $g(0)=0$;

(iii) $G_{\xi}(t)=t-g^{*}(\xi)$ for each $t\in\mathbb{R}$ and $\xi\in\Delta$.

A.2.3 Topological Representation

Next we give a topological version of Theorem 36. Also in this case we only consider the monotone case, and we refer to [7] for the general case and for a proof.

Theorem 41 *A function $g : X \to \mathbb{R}$ is uniformly continuous, quasiconcave, and monotone if and only if $g(x) = \min_{\xi \in \Delta} G_\xi(\langle \xi, x \rangle)$ for all $x \in X$, $\operatorname{dom}(G_\xi) \in \{\emptyset, \mathbb{R}\}$ for all $\xi \in \Delta$, and $\{G_\xi\}_{\xi \in \Delta : \operatorname{dom}(G_\xi) = \mathbb{R}}$ are uniformly equicontinuous.*[27]

A.2.4 Uniqueness

Theorem 5 is based on the following result, proved in [7].

Lemma 42 *Suppose $g : X \to [-\infty, \infty]$ and $G : \mathbb{R} \times \Delta \to [-\infty, \infty]$ satisfy the following conditions:*

(i) $\lim_{t \to \infty} G(t, \xi) = \lim_{t \to \infty} G(t, \xi')$ *for all* $\xi, \xi' \in \Delta$;

(ii) $G(\cdot, \xi)$ *is increasing for each* $\xi \in \Delta$;

(iii) G *is lower semicontinuous and quasiconvex on* $\mathbb{R} \times \Delta$;

(iv) $g(x) = \inf_{\xi \in \Delta} G(\langle \xi, x \rangle, \xi)$ *for all* $x \in X$.

Then, $G(t, \xi) = \sup_{x \in X : \langle \xi, x \rangle \leq t} g(x) = G_\xi(t)$ *for all* $(t, \xi) \in \mathbb{R} \times \Delta$.

Remark 43 *Since $\xi \mapsto G(\langle \xi, x \rangle, \xi)$ is lower semicontinuous, the* inf *in (iv) is attained.*

A.2.5 Positively Homogeneous Functionals

As proved in [7], Theorem 41 takes a special form for positively homogeneous and quasiconcave functionals such that $g(x) \neq 0$ for some $x \in X_+$. Begin with a lemma.

Lemma 44 *Let $\widetilde{\Delta}$ be a nonempty, closed, and convex subset of Δ, $c_1 : \widetilde{\Delta} \to (0, \infty)$ concave and upper semicontinuous, and $c_2 : \widetilde{\Delta} \to (0, \infty]$ convex and lower semicontinuous. Let*

$$G(t, \xi) = \begin{cases} \frac{t}{c_1(\xi)} & \text{if } t \geq 0 \text{ and } \xi \in \widetilde{\Delta} \\ \frac{t}{c_2(\xi)} & \text{if } t \leq 0 \text{ and } \xi \in \widetilde{\Delta} \\ \infty & \text{if } \xi \in \Delta \setminus \widetilde{\Delta} \end{cases}$$

[27] That is, for every $\varepsilon > 0$ there is $\delta > 0$ such that $|t - t'| \leq \delta$ implies $|G_\xi(t) - G_\xi(t')| \leq \varepsilon$, for all $t, t' \in \mathbb{R}$ and all $\xi \in \Delta$ such that $\operatorname{dom}(G_\xi) = \mathbb{R}$.

and

$$g(x) = \inf_{\xi \in \Delta} G(\langle \xi, x \rangle, \xi) \qquad \forall x \in X. \tag{46}$$

Then g is finite, monotone, upper semicontinuous, positively homogeneous, quasiconcave and

$$g(x) = \min_{\xi \in \widetilde{\Delta}} \left(\frac{\langle \xi, x \rangle^+}{c_1(\xi)} - \frac{\langle \xi, x \rangle^-}{c_2(\xi)} \right) \qquad \forall x \in X. \tag{47}$$

Moreover:

- $G_\xi(t) = G(t, \xi)$ *for all* $(t, \xi) \in \mathbb{R} \times \Delta$;
- *if* $\widetilde{\Gamma}$ *is a nonempty, closed, and convex subset of* Δ, $d_1 : \widetilde{\Gamma} \to (0, \infty)$ *is concave and upper semicontinuous,* $d_2 : \widetilde{\Gamma} \to (0, \infty]$ *is convex and lower semicontinuous, and*

  $$g(x) = \min_{\xi \in \widetilde{\Gamma}} \left(\frac{\langle \xi, x \rangle^+}{d_1(\xi)} - \frac{\langle \xi, x \rangle^-}{d_2(\xi)} \right) \qquad \forall x \in X \tag{48}$$

  then $\left(\widetilde{\Gamma}, d_1, d_2\right) = \left(\widetilde{\Delta}, c_1, c_2\right)$;
- g *is non-negative and concave on* $\widetilde{\Delta}^{\oplus} = \left\{ x \in X : \langle \xi, x \rangle \geq 0 \text{ for all } \xi \in \widetilde{\Delta} \right\}$;
- g *concave on* X *if and only if* $c_1(\xi) \geq c_2(\xi)$ *for all* $\xi \in \widetilde{\Delta}$.

We can now state the announced version of Theorem 41.

Theorem 45 *Let* $g : X \to \mathbb{R}$ *be such that* $g(x) \neq 0$ *for some* $x \in X_+$. *Then g is monotone, quasiconcave, uniformly continuous, and positively homogeneous if and only if*

$$g(x) = \min_{\xi \in \Delta} G_\xi(\langle \xi, x \rangle) \qquad \forall x \in X \tag{49}$$

and there exist a nonempty, closed, and convex subset $\widetilde{\Delta}$ *of* Δ, $c_1 : \widetilde{\Delta} \to (0, \infty)$ *concave and upper semicontinuous with* $\inf_{\xi \in \widetilde{\Delta}} c_1(\xi) > 0$, *and* $c_2 : \widetilde{\Delta} \to (0, \infty]$ *convex and lower semicontinuous, such that*

$$G_\xi(t) = \begin{cases} \frac{t}{c_1(\xi)} & \text{if } t \geq 0 \text{ and } \xi \in \widetilde{\Delta} \\ \frac{t}{c_2(\xi)} & \text{if } t \leq 0 \text{ and } \xi \in \widetilde{\Delta} \\ \infty & \text{if } \xi \in \Delta \setminus \widetilde{\Delta}. \end{cases} \tag{50}$$

Moreover, g is normalized if and only if $\max_{\xi \in \widetilde{\Delta}} c_1(\xi) = \min_{\xi \in \widetilde{\Delta}} c_2(\xi) = 1$.

A.3 Continuity of Monotone Functionals

A.3.1 Lower and Upper Continuity

Lemma 46 *Let Y be lower open and convex. For a monotone function $g : Y \to \mathbb{R}$ the following conditions are equivalent:*

(i) g is left continuous;

(ii) g is lower semicontinuous;

(iii) for any $c \in \mathbb{R}$ and $x, y \in Y$, the set $\{\alpha \in [0,1] : g(\alpha x + (1-\alpha) y) \leq c\}$ is closed;

(iv) for any $c \in \mathbb{R}$ and $x, y \in Y$ with $y \leq x$ and $g(x) > c$, there is $\alpha \in (0,1)$ such that $g(\alpha x + (1-\alpha) y) > c$.

Proof. (i) implies (ii). Let $c \in \mathbb{R}$, $S(g,c) = \{x \in Y : g(x) \leq c\}$. We want to show that, $\{x_n\}_{n\in\mathbb{N}} \subseteq S(g,c)$ and $x_n \to x \in Y$ imply $x \in S(g,c)$. There is $\varepsilon_0 > 0$ such that $x - \varepsilon e \in Y$ for all $\varepsilon \in [0, \varepsilon_0]$. Let $\varepsilon_m > 0$ be such that $\{\varepsilon_m\}_{m\in\mathbb{N}} \in [0, \varepsilon_0]$ and $\varepsilon_m \downarrow 0$. Then $x - \varepsilon_m e \in Y$ for all $m \in \mathbb{N}$. Since $x_n \to x$, for all $m \in \mathbb{N}$ there is $n_m \in \mathbb{N}$ such that $x - \varepsilon_m e \leq x_{n_m}$ and monotonicity implies $g(x - \varepsilon_m e) \leq g(x_{n_m}) \leq c$, and left continuity guarantees $g(x) = \lim_m g(x - \varepsilon_m e) \leq c$. (This implication does not require convexity.)

(ii) implies (iii). Let $c \in \mathbb{R}$ and $x, y \in Y$. Since Y is convex, $\alpha x + (1-\alpha) y \in Y$ for all $\alpha \in [0,1]$. Let $\{\alpha_n\}_{n\in\mathbb{N}} \subseteq [0,1]$ be such that $\alpha_n \to \alpha_0$ and $g(\alpha_n x + (1-\alpha_n) y) \leq c$. Then $\alpha_n x + (1-\alpha_n) y \in S(g,c)$ and $\alpha_n x + (1-\alpha_n) y \to \alpha_0 x + (1-\alpha_0) y \in Y$, lower semicontinuity implies $\alpha_0 x + (1-\alpha_0) y \in S(g,c)$ (i.e. $g(\alpha_0 x + (1-\alpha_0) y) \leq c$). (This implication does not require lower openness.)

(iii) implies (iv). Let $c \in \mathbb{R}$ and $x, y \in Y$ (with $y \leq x$) and $g(x) > c$. Assume, per contra, $g(\alpha x + (1-\alpha) y) \leq c$ for all $\alpha \in (0,1)$. By (iii) the set $A = \{\alpha \in [0,1] : g(\alpha x + (1-\alpha) y) \leq c\}$ is closed, thus $(0,1) \subseteq A$ implies $[0,1] = A$ and (for $\alpha = 1$) we have $g(x) \leq c$, which is absurd. (This implication does not require lower openness.)

(iv) implies (i). Let $x_n \nearrow x_0$ in Y. Monotonicity guarantees $g(x_n) \uparrow c \leq g(x_0)$. Assume, per contra, $g(x_0) > c$. By (iv), for each $y \in Y$ with $y \leq x_0$, there is $\alpha_y \in (0,1)$ such that $g((1-\alpha_y) x_0 + \alpha_y y) > c$. Take $\varepsilon_0 > 0$ such that $x_0 - \varepsilon_0 e \in Y$. Set $y = x_0 - \varepsilon_0 e$ and notice that $Y \ni (1-\alpha_y) x_0 + \alpha_y y = x_0 - \alpha_y x_0 + \alpha_y x_0 - \alpha_y \varepsilon_0 e = x_0 - \alpha_y \varepsilon_0 e$. Since $x_n \to x_0$, there is $\bar{n} \in N$ such that $x_n \geq x_0 - \alpha_y \varepsilon e = (1-\alpha_y) x_0 + \alpha_y y$ for all $n \geq \bar{n}$ and $g(x_n) \geq g((1-\alpha_y) x_0 + \alpha_y y) > c$, which is absurd. ■

If X is hyper-Archimedean, and Y is replaced by a (non-necessarily lower open) set of the form $X(T)$ the above results still hold; more indeed is true:

Proposition 47 *Let X be hyper-Archimedean. For a monotone function $g : X(T) \to \mathbb{R}$, conditions (i)-(iv) of Lemma 46 are equivalent. Moreover, lower semicontinuity is also equivalent to the following conditions:*

(v) for any $k \in T$, $c \in \mathbb{R}$ and $x \in X(T)$, the set $\{\alpha \in [0,1] : g(\alpha x + (1-\alpha) ke) \leq c\}$ is closed;

(vi) for any $k \in T$, $c \in \mathbb{R}$ and $x \in X(T)$ with $g(x) > c$, there is $\alpha \in (0,1)$ such that $g(\alpha x + (1-\alpha) ke) > c$.

(vii) for any $k \in T$, $c \in \mathbb{R}$ and $x \in X(T)$ with $ke \leq x$ and $g(x) > c$, there is $\alpha \in (0,1)$ such that $g(\alpha x + (1-\alpha) ke) > c$.

Lemma 48 *Let X be hyper-Archimedean. If $x_n, x_0 \in X(T)$, $x_n \to x_0$, and $\operatorname{ess\,inf}(x_0) = \inf T$, then for all $\alpha \in (0,1)$ there is $\bar{n} = \bar{n}_\alpha \in \mathbb{N}$ such that $x_n \geq \alpha x_0 + (1-\alpha)(\inf T) e$ for all $n \geq \bar{n}$.*

Proof. Wlog, $X = B_0(S, \Sigma)$ and $e = 1_S$. The condition $\operatorname{ess\,inf}(x_0) = \inf T$ implies $\inf T \in T$.

Let $\inf T = 0$. There exists a partition $\{A_0, A_1, ..., A_m\}$ of S in Σ and $0 = \beta_0 < \beta_1 < ... < \beta_m$ such that $x_0 = \sum_{i=0}^{m} \beta_i 1_{A_i}$. Take $\varepsilon = \min_{i=1,...,m} \beta_i - \alpha\beta_i > 0$. Since $x_n \to x_0$ there exists $\bar{n} \in \mathbb{N}$ such that

$$x_0 - \varepsilon e \leq x_n \leq x_0 + \varepsilon e \qquad \forall n \geq \bar{n}.$$

In particular, for all $n \geq \bar{n}$, if $s \in A_0$, $\alpha x_0(s) = 0 \leq x_n(s)$, else there is $i \in \{1, ..., m\}$ such that $s \in A_i$ and

$$x_n(s) \geq x_0(s) - \varepsilon \geq \beta_i + \alpha\beta_i - \beta_i = \alpha\beta_i = \alpha x_0(s)$$

and $x_n \geq \alpha x_0$, as wanted.

Let $\inf T = t$, then $x_n - te, x_0 - te \in X(T - t)$, $x_n - te \to x_0 - te$, and $\operatorname{ess\,inf}(x_0 - te) = \operatorname{ess\,inf}(x_0) - t = 0 = \inf(T - t)$. By what we have just shown, for all $\alpha \in (0,1)$ there exists $\bar{n} \in \mathbb{N}$ such that $x_n - te \geq \alpha(x_0 - te) + (1-\alpha)(\inf T - t) e = \alpha x_0 + (1-\alpha)(\inf T) e - te$ and $x_n \geq \alpha x_0 + (1-\alpha)(\inf T) e$ for all $n \geq \bar{n}$. ■

Proof of Proposition 47. If T is lower open, then $X(T)$ is lower open too, and Lemma 46 delivers the equivalence of (i)-(iv). Assume $t = \inf T \in T$.

(i) implies (ii). Let $c \in \mathbb{R}$, $S(g,c) = \{x \in X(T) : g(x) \leq c\}$. We want to show that, $\{x_n\}_{n\in\mathbb{N}} \subseteq S(g,c)$ and $x_n \to x \in X(T)$ imply $x \in S(g,c)$. Let $\varepsilon_m > 0$ be such that $\varepsilon_m \downarrow 0$ and set $y_m = (x - \varepsilon_m e) \vee te$ for all $m \in \mathbb{N}$. $\{y_m\}_{m\in\mathbb{N}} \subseteq X(T)$ and $y_m \nearrow x$. In fact, $T \ni t \leq \operatorname{ess\,inf}\left((x-\varepsilon_m e) \vee te\right) \leq \operatorname{ess\,sup}\left(((x-\varepsilon_m e) \vee te)\right) = \operatorname{ess\,sup}\left((x-\varepsilon_m e)\right) \vee t \leq \operatorname{ess\,sup}(x) \in T$, moreover, $(x-\varepsilon_m e) \leq (x-\varepsilon_{m+1} e)$ implies $(x-\varepsilon_m e) \leq (x-\varepsilon_{m+1} e) \vee te$ and $(x-\varepsilon_m e) \vee te \leq (x-\varepsilon_{m+1} e) \vee te$, thus y_m is increasing and $x-\varepsilon_m e \leq (x-\varepsilon_m e) \vee te \leq x$ implies $y_m \to x$. Since $x_n \to x$, for all $m \in \mathbb{N}$ there is $n_m \in \mathbb{N}$ such that $x-\varepsilon_m e \leq x_{n_m}$ and $x_{n_m} \in X(T)$ implies $te \leq x_{n_m}$, whence $y_m \leq x_{n_m}$ and $g(y_m) \leq g(x_{n_m}) \leq c$, left continuity guarantees $g(x) = \lim_m g(y_m) \leq c$.[28]

(ii) implies (iii) and (iii) implies (iv) are proved in exactly the same way as in Lemma 46.

(iv) implies (i). Let $x_n \nearrow x_0$ in $X(T)$. Monotonicity guarantees $g(x_n) \uparrow c \leq g(x_0)$. Assume, per contra, $g(x_0) > c$. By (iv), for each $y \in X(T)$ with $y \leq x_0$, there is $\alpha_y \in (0,1)$ such that $g((1-\alpha_y)x_0 + \alpha_y y) > c$. If $\operatorname{ess\,inf}(x_0) > \inf T$, there is $\varepsilon_0 > 0$ such that $x_0 - \varepsilon_0 e \in X(T)$. Set $y = x_0 - \varepsilon_0 e$ and notice that $X(T) \ni (1-\alpha_y)x_0 + \alpha_y y = x_0 - \alpha_y x_0 + \alpha_y x_0 - \alpha_y \varepsilon_0 e = x_0 - \alpha_y \varepsilon_0 e$. Since $x_n \to x_0$, there is $\bar{n} \in N$ such that $x_n \geq x_0 - \alpha_y \varepsilon e = (1-\alpha_y)x_0 + \alpha_y y$ for all $n \geq \bar{n}$ and $g(x_n) \geq g((1-\alpha_y)x_0 + \alpha_y y) > c$, which is absurd.

Else if $\operatorname{ess\,inf}(x_0) = \inf T = t$. Set $y = te$, by Lemma 48, there is $\bar{n} = \bar{n}_{\alpha_y} \in \mathbb{N}$ such that $x_n \geq \alpha_y x_0 + (1-\alpha)y$ for all $n \geq \bar{n}$, and $g(x_n) \geq g((1-\alpha_y)x_0 + \alpha_y y) > c$, which is absurd.

We have shown that (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) $\Rightarrow$ (iv) $\Rightarrow$ (i).

Clearly (iii) implies (v), the proof of (v) implies (vi) is almost identical to the one of (iii) implies (iv), and obviously, (vi) implies (vii). It only remains to show that (vii) implies (i), which is almost identical to (iv) $\Rightarrow$ (i):

(vii) implies (i). Let $x_n \nearrow x_0$ in $X(T)$. Monotonicity guarantees $g(x_n) \uparrow c \leq g(x_0)$. Assume, per contra, $g(x_0) > c$. By (vii), for each $k \in T$ with $ke \leq x_0$, there is $\alpha_k \in (0,1)$ such that $g((1-\alpha_k)x_0 + \alpha_k ke) > c$. If $\operatorname{ess\,inf}(x_0) > \inf T$, choose $k \in T$ such that $\operatorname{ess\,inf}(x_0) > k > \inf T$, and set $\varepsilon = \operatorname{ess\,inf}(x_0) - k > 0$. Then $x_0 - \varepsilon e \in X(T)$ and $x_0 - \alpha_k \varepsilon e \in X(T)$ too. In fact, $\inf T < k = \operatorname{ess\,inf}(x_0) - \varepsilon = \operatorname{ess\,inf}(x_0 - \varepsilon e) \leq \operatorname{ess\,sup}(x_0 - \varepsilon e) \leq \operatorname{ess\,sup}(x_0) \in T$. Therefore there is $\bar{n} \in N$ such that $x_n \geq x_0 - \alpha_k \varepsilon e = x_0 - \alpha_k \operatorname{ess\,inf}(x_0)e + \alpha_k ke \geq x_0 - \alpha_k x_0 + \alpha_k ke = (1-\alpha_k)x_0 + \alpha_k ke$ for all $n \geq \bar{n}$ and $g(x_n) \geq g((1-\alpha_k)x_0 + \alpha_k ke) > c$, which is absurd.

Else if $\operatorname{ess\,inf}(x_0) = \inf T = t$. Set $k = t$, by Lemma 48, there is $\bar{n} = \bar{n}_{\alpha_k} \in \mathbb{N}$ such that $x_n \geq \alpha_k x_0 + (1-\alpha_k)ke$ for all $n \geq \bar{n}$, and $g(x_n) \geq g((1-\alpha_k)x_0 + \alpha_k ke) > c$, which is absurd. ■

[28] Notice that we did not use the hyper-archimedean assumption. Therefore a monotone function $g : X(T) \to \mathbb{R}$ is left continuous if and only if it is lower semicontinuous. (For the "if" part, observe that $x_n \to x$ in $X(T)$ and $x_n \leq x_{n+1}$ for all $n \in \mathbb{N}$ imply $x_n \leq x$ for all $n \in \mathbb{N}$. Monotonicity of g implies $g(x_n) \uparrow c \leq g(x)$ and lower semicontinuity delivers $c = \lim_n g(x_n) \geq g(x)$.

Very similar results hold for upper semicontinuity: just observe that $g(x)$ from $X(T)$ to $\mathbb{R}$ is lower semicontinuous and monotone if and only if $-g(-x)$ from $X(-T)$ to $\mathbb{R}$ is upper semicontinuous and monotone.

A.3.2 Uniform Continuity and Lipschitzianity

Proposition 49 *For a monotone $g : X(T) \to \mathbb{R}$ the following properties are equivalent:*

(i) g is uniformly continuous on $X(T)$;

(ii) for every $\varepsilon > 0$ there is $\delta \in (0, \sup T - \inf T)$ such that

$$g(x + \delta e) \leq g(x) + \varepsilon \tag{51}$$

for all $x \in X(T)$ with $x + \delta e \in X(T)$.

Notice that if T is bounded and $\delta > \sup T - \inf T$, then (iii) is vacuously satisfied since there is no $x \in X(T)$ such that $x + \delta e \in X(T)$.

Proof. (i) implies (ii). Fix $\varepsilon > 0$ and let $\delta' > 0$ be such that $\|x - y\| \leq \delta'$ implies $|g(x) - g(y)| \leq \varepsilon$. Set $\delta = 2^{-1} \min\{\delta', \sup T - \inf T\}$. If $x, x+\delta e \in X(T)$, then $g(x + \delta e) - g(x) \leq |g(x + \delta e) - g(x)| \leq \varepsilon$.

(ii) implies (i). Fix $\varepsilon > 0$ and let $\delta \in (0, \sup T - \inf T)$ be such that $g(x + \delta e) \leq g(x) + \varepsilon$ for all $x \in X(T)$ such that $x + \delta e \in X(T)$. Notice that if x and $x - \delta e$ belong to $X(T)$, then $(x - \delta e)$ and $(x - \delta e) + \delta e \in X(T)$. Thus, $g(x) = g((x - \delta e) + \delta e) \leq g(x - \delta e) + \varepsilon$ and $g(x - \delta e) \geq g(x) - \varepsilon$. Let $x, y \in X(T)$ be such that $\|x - y\| \leq \delta$. Then

$$x - \delta e \leq y \leq x + \delta e. \tag{52}$$

Moreover:

Claim. There exist $t, \tau \in T$ such that $t + \delta \leq \tau$ and $te \leq x, y \leq \tau e$.

Proof of the Claim. Set $t' = \operatorname{ess\,inf}(x \wedge y) = \operatorname{ess\,inf}(x) \wedge \operatorname{ess\,inf}(y) \in T$ and $\tau' = \operatorname{ess\,sup}(x \vee y) = \operatorname{ess\,sup}(x) \vee \operatorname{ess\,sup}(y) \in T$. Clearly $t' \leq \tau'$ and $t'e \leq x, y \leq \tau' e$. If $\tau' - t' \geq \delta$ set $t = t'$ and $\tau = \tau'$. Otherwise, consider the following cases: (i) if T is unbounded above, set $t = t'$ and $\tau = \tau' + \delta$; (ii) if T is unbounded below, set $t = t' - \delta$ and $\tau = \tau'$; (iii) if T is bounded consider two sequences t'_n and τ'_n in T such $t'_1 = t'$, $\tau'_1 = \tau'$, $t'_n \downarrow \inf T$, $\tau'_n \uparrow \sup T$. For all $n \geq 1$, $t'_n e \leq x, y \leq \tau'_n e$ and $\tau'_n - t'_n \to \sup T - \inf T > \delta$. Hence there is $\bar{n} \in \mathbb{N}$ such that $\tau'_{\bar{n}} - t'_{\bar{n}} > \delta$; set $t = t'_{\bar{n}}$ and $\tau = \tau'_{\bar{n}}$. $\square$

Since

$$t \le \operatorname{ess\,inf}\left((x-\delta e)\vee te\right) \le \operatorname{ess\,sup}\left((x-\delta e)\vee te\right) \le \operatorname{ess\,sup}(x) \le \tau$$
$$t \le \operatorname{ess\,inf}(x) \le \operatorname{ess\,inf}\left((x+\delta e)\wedge \tau e\right) \le \operatorname{ess\,sup}\left((x+\delta e)\wedge \tau e\right) \le \tau$$
$$t \le \operatorname{ess\,inf}(x) \le \operatorname{ess\,inf}\left(x\vee (t+\delta)e\right) \le \operatorname{ess\,sup}\left(x\vee (t+\delta)e\right) = \operatorname{ess\,sup}(x)\vee(t+\delta) \le \tau$$
$$t \le \operatorname{ess\,inf}(x)\wedge(\tau-\delta) = \operatorname{ess\,inf}\left(x\wedge(\tau-\delta)e\right) \le \operatorname{ess\,sup}\left(x\wedge(\tau-\delta)e\right) \le \operatorname{ess\,sup}(x) \le \tau$$

then $(x-\delta e)\vee te, (x+\delta e)\wedge \tau e$, $x\vee(t+\delta)e$, $x\wedge(\tau-\delta)e \in X(T)$, as well as

$$(x\vee(t+\delta)e)-\delta e = (x-\delta e)\vee te \in X(T) \text{ and } (x\wedge(\tau-\delta)e)+\delta e = (x+\delta e)\wedge\tau e \in X(T). \tag{53}$$

From (52) we have $(x-\delta e)\vee te \le y \le (x+\delta e)\wedge \tau e$. By monotonicity, (53), and the choice of δ, $g(x)-\varepsilon \le g((x\vee(t+\delta)e))-\varepsilon \le g((x\vee(t+\delta)e)-\delta e) = g((x-\delta e)\vee te) \le g(y) \le g((x+\delta e)\wedge\tau e) = g((x\wedge(\tau-\delta)e)+\delta e) \le g((x\wedge(\tau-\delta)e))+\varepsilon \le g(x)+\varepsilon$, and so $g(x)-\varepsilon \le g(y) \le g(x)+\varepsilon$, as desired. ■

A similar argument, can be used to prove the following variation:

Proposition 50 *A monotone $g : X(T) \longrightarrow \mathbb{R}$ is ℓ-Lipschitz on $X(T)$ if and only if $g(x+\delta e) \le g(x)+\ell\delta$ for all $x\in X(T)$ and all $\delta>0$ such that $x+\delta e\in X(T)$.*

A.3.3 Linear Continuity

Lemma 51 *If $G\in\mathcal{G}(T\times\Delta)$, then $g(x) = \inf_{\xi\in\Delta} G(\langle\xi,x\rangle,\xi)$ for all $x\in X(T)$ is finite, (evenly) quasiconcave, monotone, normalized, and $G(t,\xi)\ge G_\xi(t)$ for all $(t,\xi)\in T\times\Delta$.*

Proof. We only assume $G : T\times\Delta \to (-\infty,\infty]$ is increasing in the first component and $\inf_{p\in\Delta} G(t,p) = t$ for all $t\in T$.

We first prove monotonicity: if $x\ge y$, then $\langle\xi,x\rangle \ge \langle\xi,y\rangle$ for all $\xi\in\Delta$, and monotonicity of $G(\cdot,\xi)$ implies that $G(\langle\xi,x\rangle,\xi) \ge G(\langle\xi,y\rangle,\xi)$, and hence $g(x)\ge g(y)$.

As to normalization: $g(te) = \inf_{\xi\in\Delta} G(\langle\xi,te\rangle,\xi) = \inf_{\xi\in\Delta} G(t,\xi) = t$ for all $t\in T$.

Finiteness follows from monotonicity and normalization, in fact, for all $x\in X(T)$, $\operatorname{ess\,inf}(x)e \le x \le \operatorname{ess\,sup}(x)e$ implies $\operatorname{ess\,inf}(x) \le g(x) \le \operatorname{ess\,sup}(x)$.

Next we show (even) quasiconcavity: Let $\alpha\in\mathbb{R}$. As observed, $X(T)$ is evenly quasiconvex, thus the set

$$L = X(T) \cap \bigcap_{(\xi,b)\in\Delta\times\mathbb{R}:[\xi>b]\supseteq\{y\in X(T):g(y)\ge a\}} [\xi>b]$$

is evenly quasiconvex and contains $\{y\in X(T) : g(y)\ge a\}$.

Let $\bar{x} \notin \{y \in X(T) : g(y) \geq a\}$. Then, either $\bar{x} \notin X(T)$ and so $\bar{x} \notin L$; or $\bar{x} \in X(T)$ and $a > g(\bar{x}) = \inf_{\xi \in \Delta} G(\langle \xi, \bar{x} \rangle, \xi)$, then there is $\bar{\xi} \in \Delta$ such that $G(\langle \bar{\xi}, \bar{x} \rangle, \bar{\xi}) < a$, and $g(y) \leq G(\langle \bar{\xi}, y \rangle, \bar{\xi}) \leq G(\langle \bar{\xi}, \bar{x} \rangle, \bar{\xi}) < a$ for all $y \in X(T)$ such that $\bar{\xi}(y) \leq \bar{\xi}(\bar{x})$, that is $\{y \in X(T) : \bar{\xi}(y) \leq \bar{\xi}(\bar{x})\} \subseteq \{y \in X(T) : g(y) < a\}$ and $\{y \in X(T) : g(y) \geq a\} \subseteq \{y \in X(T) : \bar{\xi}(y) > \bar{\xi}(\bar{x})\} \subseteq [\bar{\xi} > \bar{\xi}(\bar{x})]$. Thus

$$(\bar{\xi}, \bar{\xi}(\bar{x})) \in \Delta \times \mathbb{R} : [\bar{\xi} > \bar{\xi}(\bar{x})] \supseteq \{y \in X(T) : g(y) \geq a\}.$$

But, $\bar{x} \notin [\bar{\xi} > \bar{\xi}(\bar{x})]$, and so $\bar{x} \notin L$. Therefore, L is contained $\{y \in X(T) : g(y) \geq a\}$, and the two sets coincide.

Moreover, for all $(\bar{t}, \bar{\xi}) \in T \times \Delta$, and all $y \in X(T)$ such that $\langle \bar{\xi}, y \rangle \leq \bar{t}$, then $g(y) = \inf_{\xi \in \Delta} G(\langle \xi, y \rangle, \xi) \leq G(\langle \bar{\xi}, y \rangle, \bar{\xi}) \leq G(\bar{t}, \bar{\xi})$. Therefore, $G_{\bar{\xi}}(\bar{t}) = \sup_{y \in X(T) : \langle \bar{\xi}, y \rangle \leq \bar{t}} g(y) \leq G(\bar{t}, \bar{\xi})$. ■

Lemma 52 *If $G \in \mathcal{H}(T \times \Delta)$, then $g : X(T) \to \mathbb{R}$ defined by $g(x) = \inf_{\xi \in \Delta} G(\langle \xi, x \rangle, \xi)$ is continuous and the* inf *is attained for all $x \in X(T)$.*

Proof. The proof is divided into several claims that are used in different parts of the paper.

Let $G : T \times \Delta \to [-\infty, \infty]$ be lower semicontinuous. Define $\Gamma : X(T) \times \Delta \to [-\infty, \infty]$ by $\Gamma(x, \xi) = G(\langle \xi, x \rangle, \xi)$ for all $(x, \xi) \in X(T) \times \Delta$.

Claim 1. Γ is lower semicontinuous on $X(T) \times \Delta$.

Proof of Claim 1. Consider a net $\{(x_\alpha, \xi_\alpha)\}$ in $X(T) \times \Delta$ such that $(x_\alpha, \xi_\alpha) \to (x, \xi) \in X(T) \times \Delta$. This is equivalent to $x_\alpha \to x$ and $\xi_\alpha \to \xi$. It follows that $\langle \xi_\alpha, x_\alpha \rangle \to \langle \xi, x \rangle$. In fact, $|\langle \xi_\alpha, x_\alpha \rangle - \langle \xi, x \rangle| \leq |\langle \xi_\alpha, x_\alpha \rangle - \langle \xi_\alpha, x \rangle| + |\langle \xi_\alpha, x \rangle - \langle \xi, x \rangle| = |\langle \xi_\alpha, x_\alpha - x \rangle| + |\langle \xi_\alpha, x \rangle - \langle \xi, x \rangle| \leq \|\xi_\alpha\| \|x_\alpha - x\| + |\langle \xi_\alpha, x \rangle - \langle \xi, x \rangle| = \|x_\alpha - x\| + |\langle \xi_\alpha, x \rangle - \langle \xi, x \rangle| \to 0$. Since G is lower semicontinuous, it then follows that $\liminf_\alpha \Gamma(x_\alpha, \xi_\alpha) = \liminf_\alpha G(\langle \xi_\alpha, x_\alpha \rangle, \xi_\alpha) \geq G(\langle \xi, x \rangle, \xi) = \Gamma(x, \xi)$, as wanted. □

In particular, $\Gamma(x, \cdot) : \Delta \to [-\infty, \infty]$ is lower semicontinuous on Δ for all $x \in X(T)$, thus

$$g(x) = \inf_{\xi \in \Delta} \Gamma(x, \xi) = \min_{\xi \in \Delta} \Gamma(x, \xi) = \min_{\xi \in \Delta} G(\langle \xi, x \rangle, \xi) \tag{54}$$

that is the inf is attained.

Claim 2. g is lower semicontinuous on $X(T)$.

Proof of Claim 2. Consider a sequence $\{x_n\}$ in $X(T)$ such that $x_n \to x \in X(T)$. Then, there exists a subsequence $\{x_{n_k}\}$ such that $\liminf_n g(x_n) = \lim_k g(x_{n_k})$. Furthermore, by (54), for each k there exists $\xi_{n_k} \in \Delta$ such that $g(x_{n_k}) = \Gamma(x_{n_k}, \xi_{n_k})$. Since Δ is compact, there exists a subnet $\{\xi_{n_{k_\alpha}}\}$ such that $\xi_{n_{k_\alpha}} \to \bar{\xi} \in \Delta$. By Claim 1, $\liminf_n g(x_n) = \lim_k g(x_{n_k}) = \lim_\alpha g(x_{n_{k_\alpha}}) = \lim_\alpha \Gamma(x_{n_{k_\alpha}}, \xi_{n_{k_\alpha}}) \geq \Gamma(x, \bar{\xi}) \geq \min_{\xi \in \Delta} \Gamma(x, \xi) = g(x)$, as wanted. □

Now assume $G \in \mathcal{H}(T \times \Delta)$, since $G(\cdot, \xi)$ is extended-valued continuous on T for each $\xi \in \Delta$, then it is upper semicontinuous on T for each $\xi \in \Delta$. Therefore $\Gamma(\cdot, \xi) : X(T) \to [-\infty, \infty]$ is upper semicontinuous on $X(T)$ for all $\xi \in \Delta$,[29] finally $g(\cdot) = \inf_{\xi \in \Delta} \Gamma(\cdot, \xi)$ is upper semicontinuous too. ■

Lemma 53 *If $G \in \mathcal{E}(T \times \Delta)$, then $g(x) = \inf_{\xi \in \Delta} G(\langle \xi, x \rangle, \xi)$ for all $x \in X(T)$ is uniformly continuous.*

Proof. By definition, given $\varepsilon > 0$, there is $\delta > 0$ such that $|G(t, \xi) - G(t', \xi)| \leq \varepsilon$ for all $\xi \in \Delta$ with $\mathrm{dom}(G(\cdot, \xi)) = T$, and all $t, t' \in \mathbb{R}$ with $|t - t'| \leq \delta$.

Take $x, y \in X(T)$ such that $\|x - y\| \leq \delta$. Since $g(x) \in \mathbb{R}$ (see Lemma 51), there is $\xi_x \in \Delta$ such that $g(x) \geq G(\langle \xi_x, x \rangle, \xi_x) - \varepsilon$, and it must be the case that $\mathrm{dom}(G(\cdot, \xi_x)) = T$. Moreover, since $\|x - y\| \leq \delta$, then $|\langle \xi_x, x \rangle - \langle \xi_x, y \rangle| \leq \|\xi_x\| \|x - y\| \leq \delta$. By uniform equicontinuity $|G(\langle \xi_x, x \rangle, \xi_x) - G(\langle \xi_x, y \rangle, \xi_x)| \leq \varepsilon$, and so $G(\langle \xi_x, y \rangle, \xi_x) - G(\langle \xi_x, x \rangle, \xi_x) \leq \varepsilon$ thus $g(x) \geq G(\langle \xi_x, x \rangle, \xi_x) - \varepsilon \geq G(\langle \xi_x, y \rangle, \xi_x) - 2\varepsilon \geq \inf_{\xi \in \Delta} G(\langle \xi, y \rangle, \xi) - 2\varepsilon = g(y) - 2\varepsilon$. Exchanging the roles of x and y, we get $|g(x) - g(y)| \leq 2\varepsilon$ for all $x, y \in X(T)$ such that $\|x - y\| \leq \delta$, and so g is uniformly continuous. ■

A.3.4 Monotone Continuity on Function Spaces

Theorem 54 *Let Σ be a σ-algebra, and $I : B_0(\Sigma) \to \mathbb{R}$ be such that*

$$I(\varphi) = \inf_{p \in \Delta} G\left(\int \varphi dp, p\right) \tag{55}$$

where $G : \mathbb{R} \times \Delta \to (-\infty, \infty]$ is jointly lower semicontinuous, grounded,[30] and increasing in the first component. The following conditions are equivalent:

(i) I is monotone continuous (i.e., $I(\varphi_n) \uparrow I(\varphi)$ if $\varphi_n \uparrow \varphi$);

(ii) if $\varphi, \psi \in B_0(\Sigma)$, $k \in \mathbb{R}$, and $\Sigma \ni E_n \downarrow \emptyset$, then $I(\psi) > I(\varphi)$ implies that there exists $n \in \mathbb{N}$ such that $I\left(k 1_{E_n} + \psi 1_{E_n^c}\right) > I(\varphi)$;

(iii) $G(\cdot, p) \equiv \infty$ for all $p \notin \Delta^\sigma$;

(iv) there is $q \in \Delta^\sigma$ such that $\{p \in \Delta : G(t, p) \leq \alpha\}$ is a weakly compact subset of $\Delta^\sigma(q)$ for all $t, \alpha \in \mathbb{R}$.

(v) there is $q \in \Delta^\sigma$ such that $G(\cdot, p) \equiv \infty$ for all $p \notin \Delta^\sigma(q)$;

---

[29] Let $\xi \in \Delta$, if $\{x_n\}$ in $X(T)$ and $x_n \to x \in X(T)$, then $\langle \xi, x_n \rangle \to \langle \xi, x \rangle$ and $\limsup_n G(\langle \xi, x_n \rangle, \xi) \leq G(\langle \xi, x \rangle, \xi)$.

[30] That is, such that $\inf_{p \in \Delta} G(t, p) = t$ for all $t \in \mathbb{R}$.

Proof. We will use the following claim.

Claim. Let P be a subset of Δ. The following statements are equivalent:

(a) $G(\cdot,p) \equiv \infty$ for all $p \notin P$;

(b) $\bigcup_{t,\alpha\in\mathbb{R}} \{p \in \Delta : G(t,p) \leq \alpha\} \subseteq P$;

(c) $\bigcup_{m,n=1}^{\infty} \{p \in \Delta : G(m,p) \leq n\} \subseteq P$.

Proof of the Claim. If there exists $\bar{p} \notin P$ such that $\bar{p} \in \bigcup_{t,\alpha\in\mathbb{R}} \{p \in \Delta : G(t,p) \leq \alpha\}$, then $G(\bar{t},\bar{p}) \leq \bar{\alpha}$ for some $\bar{t},\bar{\alpha} \in \mathbb{R}$ and $G(\cdot,\bar{p}) \not\equiv \infty$. That is not (b) implies not (a), and (a) implies (b).

Clearly (b) implies (c).

If there exists $\bar{p} \notin P$ such that $G(\cdot,\bar{p}) \not\equiv \infty$, then there is $\bar{t} \in \mathbb{R}$ such that $G(\bar{t},\bar{p}) < \infty$, therefore there is $\bar{n} \in \mathbb{N}$ such that $G(\bar{t},\bar{p}) \leq \bar{n}$ and, by monotonicity of $G(\cdot,\bar{p})$, for all $\bar{m} \leq \bar{t}$, $G(\bar{m},\bar{p}) \leq \bar{n}$, thus exists $\bar{p} \notin P$ such that $\bar{p} \in \bigcup_{m,n=1}^{\infty} \{p \in \Delta : G(m,p) \leq n\}$. That is not (a) implies not (c), and (c) implies (a). $\square$

(i) implies (ii). Suppose first that $k \leq \min \psi$. Set $\psi_n = k1_{E_n} + \psi 1_{E_n^c}$, then $\psi_n \uparrow \psi$ and $I(\psi_n) \uparrow I(\psi)$. Therefore there is n_0 such that $I(\psi_{n_0}) > I(\varphi)$. If $k > \min \psi$, then $k1_{E_n} + \psi 1_{E_n^c} \geq (\min \psi) 1_{E_n} + \psi 1_{E_n^c}$, but there is n_0 such that $I\left((\min \psi) 1_{E_{n_0}} + \psi 1_{E_{n_0}^c}\right) > I(\varphi)$, by monotonicity $I\left(k1_{E_{n_0}} + \psi 1_{E_{n_0}^c}\right) \geq I\left((\min \psi) 1_{E_{n_0}} + \psi 1_{E_{n_0}^c}\right) > I(\varphi)$.

(ii) implies (iii). By the Claim, it is enough to show that $\{p \in \Delta : G(t,p) \leq \alpha\} \subseteq \Delta^\sigma$ for all t and α in $\mathbb{R}$. Let $E_n \downarrow \emptyset$ and $r \in \{p \in \Delta : G(t,p) \leq \alpha\}$. Set $\varphi \equiv \alpha$ and $\psi \equiv \beta$ with $\beta > \alpha \vee 0$. For each $k > 0$ there is $n_k \geq 1$ such that $\alpha = I(\varphi) < I\left(-k1_{E_{n_k}} + \beta 1_{E_{n_k}^c}\right)$.[31] Thus, since $-k1_{E_{n_k}} + \beta 1_{E_{n_k}^c} \leq -k1_{E_n} + \beta 1_{E_n^c}$ for all $n \geq n_k$, it follows $\alpha < I\left(-k1_{E_n} + \beta 1_{E_n^c}\right) = \inf_{p\in\Delta} G\left(\left\langle p, -k1_{E_n} + \beta 1_{E_n^c}\right\rangle, p\right)$ for all $n \geq n_k$. If $\left\langle r, -k1_{E_n} + \beta 1_{E_n^c}\right\rangle \leq t$ for some $n \geq n_k$, monotonicity of $G(\cdot,r)$ implies $\inf_{p\in\Delta} G\left(\left\langle p, -k1_{E_n} + \beta 1_{E_n^c}\right\rangle, p\right) \leq G\left(\left\langle r, -k1_{E_n} + \beta 1_{E_n^c}\right\rangle, r\right)$ $\leq G(t,r) \leq \alpha$, which is absurd. Conclude that $\left\langle r, -k1_{E_n} + \beta 1_{E_n^c}\right\rangle > t$ for all $n \geq n_k$ hence $r(E_n) < k^{-1}(\beta - t)$ and so $\lim_n r(E_n) \leq k^{-1}(\beta - t)$ for each $k > 0$, finally $\lim_n r(E_n) = 0$, i.e., $r \in \Delta^\sigma$.

(iii) implies (iv). By (iii) and the Claim, $\{p \in \Delta : G(t,p) \leq \alpha\} \subseteq \Delta^\sigma$ for all $t, \alpha \in \mathbb{R}$, moreover it is weak* compact (by lower semicontinuity of G), and so, being included in Δ^σ, weakly compact (e.g., [24, Prop. 2.13]). Then, for all $m, n \in \mathbb{N}$, there is $q_{(n,m)} \in \Delta^\sigma$ such

[31] The first equality descends from the normalization of I that corresponds to the groundedness of G.

that $p \ll q_{(n,m)}$ whenever $p \in \Delta$ and $G(m,p) \leq n$. Given an enumeration $h : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ of $\mathbb{N} \times \mathbb{N}$, set $q = \sum_{(n,m)\in\mathbb{N}\times\mathbb{N}} 2^{-h(n,m)} q_{(n,m)}$. Then, $\bigcup_{m,n=1}^{\infty} \{p \in \Delta : G(m,p) \leq n\} \subseteq \Delta^{\sigma}(q)$. Let $t, \alpha \in \mathbb{R}$, and choose $m < t$ and $n \geq \alpha$, then $G(t,p) \leq \alpha$ and monotonicity of $G(\cdot, p)$ implies $G(m,p) \leq G(t,p) \leq \alpha \leq n$, that is $\{p \in \Delta : G(t,p) \leq \alpha\} \subseteq \{p \in \Delta : G(m,p) \leq n\} \subseteq \Delta^{\sigma}(q)$.

(iv) implies (v) descends immediately from the claim.

(v) implies (i). Let $\varphi_n \uparrow \varphi_0$. For each $n \geq 0$, define $\gamma_n : \Delta \to (-\infty, +\infty]$ by $\gamma_n(p) = G\left(\int \varphi_n dp, p\right)$. Each γ_n is weak* lower semicontinuous, and the sequence $\{\gamma_n\}$ is increasing. Moreover, γ_n pointwise converges to γ_0, i.e., $\gamma_n \uparrow \gamma_0$. For, if $p \in \Delta^{\sigma}(q)$, then $\int \varphi_n dp \uparrow \int \varphi_0 dp$ by the Levi Monotone Converge Theorem (notice that φ_1 is bounded below), and so, since $G(\cdot, p)$ is lower semicontinuous and increasing on $\mathbb{R}$, $\lim_n G\left(\int \varphi_n dp, p\right) = G\left(\int \varphi_0 dp, p\right)$. If $p \notin \Delta^{\sigma}(q)$, then $\gamma_n(p) = \infty$ for all $n \in \mathbb{N}$.

We conclude that γ_n pointwise converges (and so, by [14, Rem. 5.5], Γ-converges) to γ_0. By [14, Thm. 7.4], $\min_{p\in\Delta} \gamma_n(p) \to \min_{p\in\Delta} \gamma_0(p)$, that is $I(\varphi_n) \to I(\varphi_0)$, and monotonicity of I delivers: $I(\varphi_n) \uparrow I(\varphi_0)$. ■

B Integrals which are Concave Functionals

Let $\phi : \mathbb{R} \to \mathbb{R}$ be an increasing and concave function. Motivated by the study of smooth preferences, we are interested in the concave functionals $g : X \to \mathbb{R}$ given by

$$g(x) = \int_{\Delta} \phi(\langle \xi, x \rangle)\, d\mu(\xi) \qquad \forall x \in X \tag{56}$$

where μ is a countably additive Borel probability measure on the simplex Δ, i.e. $\mu \in \Delta^{\sigma}(\mathcal{B}(\Delta))$.

To study the functional (56) we need some notation: $ca(\mathcal{B}(\Delta))$ is the set of all countably additive elements of $ba(\mathcal{B}(\Delta))$, $ca_+(\mathcal{B}(\Delta)) = ca(\mathcal{B}(\Delta)) \cap ba_+(\mathcal{B}(\Delta))$ is its positive cone. Moreover, $ba(\mathcal{B}(\Delta), \mu) = \{\nu \in ba(\mathcal{B}(\Delta)) : B \in \mathcal{B}(\Delta) \text{ and } \mu(B) = 0 \text{ implies } \nu(B) = 0\}$ is (isometrically isomorphic to, e.g., [55, Ch. IV.9]) the dual of $L^{\infty}(\mu) = L^{\infty}(\Delta, \mathcal{B}(\Delta), \mu)$ and $ca(\mathcal{B}(\Delta), \mu) = ca(\mathcal{B}(\Delta)) \cap ba(\mathcal{B}(\Delta), \mu)$ is (isometrically isomorphic to) $L^1(\mu)$ (via the Radon-Nikodym derivative $\nu \mapsto d\nu/d\mu$).

Consider the mapping $A : X \to L^{\infty}(\mu)$ defined by $Ax = \langle \cdot, x \rangle$ for all $x \in X$. A is well defined since $\langle \cdot, x \rangle$ is affine and continuous on the compact set Δ, then it belongs to $L^{\infty}(\mu)$. A is linear, in fact, for all $x, y \in X$ and $\alpha \in \mathbb{R}$, $A(\alpha x + y)(\xi) = \langle \xi, \alpha x + y \rangle = \alpha \langle \xi, x \rangle + \langle \xi, y \rangle = \alpha (Ax)(\xi) + (Ay)(\xi) = (\alpha Ax + Ay)(\xi)$ for all $\xi \in \Delta$, hence $A(\alpha x + y) = \alpha Ax + Ay$. A is bounded, in fact, $|Ax(\xi)| = |\langle \xi, x \rangle| \leq \|\xi\| \|x\| = \|x\|$ for all $\xi \in \Delta$ and $x \in X$, thus

$\|Ax\|_{L^\infty(\mu)} \leq 1\|x\|$ for all $x \in X$, and $|||A||| \leq 1$. Then A is continuous, and obviously positive.

Its adjoint is $A^* : ba(\mathcal{B}(\Delta), \mu) \to X^*$ is defined, for all $\nu \in ba(\mathcal{B}(\Delta), \mu)$ by $A^*\nu = \nu A$, that is

$$\langle A^*\nu, x\rangle = \langle \nu, Ax\rangle = \int_\Delta Ax d\nu = \int_\Delta \langle \xi, x\rangle d\nu(\xi) \qquad \forall x \in X. \tag{57}$$

A^* is continuous and $A^*\nu$ is denoted by $\int_\Delta \xi d\nu(\xi)$ in view of (57).

Moreover, A^* is obviously positive, and it preserves the norm between the positive cones $ba_+(\mathcal{B}(\Delta))$ and X_+^*. In fact, if $\nu \in ba_+(\mathcal{B}(\Delta), \mu)$, then $\|\nu\|_{ba(\mathcal{B}(\Delta),\mu)} = \nu(\Delta) = \int_\Delta 1 d\nu = \int_\Delta \langle \xi, e\rangle d\nu(\xi) = \langle A^*\nu, e\rangle = \|A^*\nu\|_{X^*}$.

For $\xi \in X_+^*$, set $\Gamma(\xi) = (A^*)^{-1}(\xi) \cap ca_+(\mathcal{B}(\Delta), \mu) = \{\nu \in ca_+(\mathcal{B}(\Delta), \mu) : A^*\nu = \xi\} = \left\{\nu \in ca_+(\mathcal{B}(\Delta), \mu) : \int_\Delta \zeta d\nu(\zeta) = \xi\right\}$. $\Gamma(\xi)$ is a (possibly empty) closed and convex (hence weakly closed) subset of $ca_+(\mathcal{B}(\Delta), \mu)$ and $\nu(\Delta) = \xi(e)$ for all $\nu \in \Gamma(\xi)$.[32] In particular, if $\xi \in \Delta$, then $\Gamma(\xi) = \left\{\nu \in \Delta^\sigma(\mu) : \int_\Delta \zeta d\nu(\zeta) = \xi\right\}$. Finally, in this case, for all $k > 0$, $\Gamma(k\xi) = k\Gamma(\xi)$, and the same is true for $k = 0$ if $\Gamma(\xi) \neq \emptyset$, while $\Gamma(k\xi) = \{0\} \neq k\Gamma(\xi) = \emptyset$ if $k = 0$ and $\Gamma(\xi) = \emptyset$. In fact,

- if $k > 0$, then $\gamma \in ca_+(\mathcal{B}(\Delta), \mu)$ and $A^*\gamma = \xi$, implies $k\gamma \in ca_+(\mathcal{B}(\Delta), \mu)$ and $A^*k\gamma = k\xi$, that is $k\Gamma(\xi) \subseteq \Gamma(k\xi)$, conversely, if $\nu \in ca_+(\mathcal{B}(\Delta), \mu)$ and $A^*\nu = k\xi$, then $\gamma = k^{-1}\nu \in ca_+(\mathcal{B}(\Delta), \mu)$ and $A^*\gamma = A^*k^{-1}\nu = k^{-1}A^*\nu = \xi$, and $\nu = k\gamma$, that is $\Gamma(k\xi) \subseteq k\Gamma(\xi)$;

- if $k = 0$, then $\nu \in ca_+(\mathcal{B}(\Delta), \mu)$ and $A^*\nu = 0$ imply $\|\nu\|_{ba(\mathcal{B}(\Delta),\mu)} = \|A^*\nu\|_{X^*} = 0$ and $\nu = 0$, that is $\Gamma(k\xi) = \{0\}$, while $k\Gamma(\xi) = \{0\}$ if $\Gamma(\xi) \neq \emptyset$ and $k\Gamma(\xi) = \emptyset$ if $\Gamma(\xi) = \emptyset$.

Theorem 55 *The functional (56) is finite, concave, continuous and monotone on X. Its conjugate is $g^*(\xi) = \sup\left\{\int_\Delta \phi^*\left(\frac{d\nu}{d\mu}(\zeta)\right) d\mu(\zeta) : \nu \in \Gamma(\xi)\right\}$ for all $\xi \in X^*$, with the convention $g^*(\xi) = -\infty$ if $\Gamma(\xi) = \emptyset$. Moreover, for all $(t, \xi) \in \mathbb{R}\times\Delta$,*

$$G_\xi(t) = \begin{cases} \inf\left\{\inf_{k\geq 0}\left[tk - \int_\Delta \phi^*\left(k\frac{d\nu}{d\mu}(\zeta)\right) d\mu(\zeta)\right] : \nu \in \Gamma(\xi)\right\} & \text{if } \Gamma(\xi) \neq \emptyset \\ \sup_{k\in\mathbb{R}} \phi(k) & \text{if } \Gamma(\xi) = \emptyset. \end{cases}$$

Proof. The properties of the functional g may be easily obtained directly but we shall get them from more general results. Our starting point is the functional $I_\phi(u) = \int_\Delta \phi(u(\xi)) d\mu(\xi)$ defined for $u \in L^\infty(\mu)$. This is a normal concave integral, studied by [47] and [48].

[32] $(A^*)^{-1}(\xi)$ is closed in $ba(\mathcal{B}(\Delta), \mu)$ since A^* is continuous, while $ca_+(\mathcal{B}(\Delta), \mu)$ is closed in $ca(\mathcal{B}(\Delta), \mu)$ which is a complete subspace of $ba(\mathcal{B}(\Delta), \mu)$. Moreover, for all $\nu \in \Gamma(\xi)$, $\nu \geq 0$ and $A^*\nu = \xi$, hence $\nu(\Delta) = \langle A^*\nu, e\rangle = \xi(e)$.

By [48, Corollary 2A], I_ϕ is finite, concave, and continuous; monotonicity immediately descends from that of ϕ. Moreover, the conjugate $I_\phi^* : ba(\mathcal{B}(\Delta), \mu) \to [-\infty, \infty)$ of I_ϕ is given by

$$I_\phi^*(\nu) = I_{\phi^*}(u^*) = \int_\Delta \phi^*(u^*(\zeta))\, d\mu(\zeta) \tag{58}$$

if there exists $u^* \in L^1(\mu)$ such that $\nu(u) = \int_\Delta u(\zeta)\, u^*(\zeta)\, d\mu(\zeta)$ for all $u \in L^\infty(\mu)$, while $I_\phi^*(\nu) = -\infty$ otherwise. By the Radon-Nikodym Theorem, the condition "there exists $u^* \in L^1(\mu)$ such that $\nu(u) = \int_\Delta u(\zeta)\, u^*(\zeta)\, d\mu(\zeta)$ for all $u \in L^\infty(\mu)$" amounts to "ν is countably additive" and in this case $u^* = d\nu/d\mu$ is unique (as an equivalence class).[33] Therefore,

$$I_\phi^*(\nu) = \begin{cases} \int_\Delta \phi^*\left(\frac{d\nu}{d\mu}(\zeta)\right) d\mu(\zeta) & \text{if } \nu \text{ is countably additive} \\ -\infty & \text{otherwise.} \end{cases} \tag{59}$$

Consider the bounded linear operator $A : x \mapsto \langle \cdot, x \rangle$ that we studied above. Clearly $g = I_\phi \circ A$ or, according to standard convex analysis notation $g = I_\phi A$. In particular, g is finite, concave, continuous, and monotone.

Since I_ϕ is finite and continuous on $L^\infty(\mu)$, [48, Theorem 3] guarantees that $g^* = (I_\phi A)^* = A^* I_\phi^*$ where A^* is the adjoint of A, and $A^* I_\phi^*$ is defined, for all $\xi \in X^*$, by

$$A^* I_\phi^*(\xi) = \sup\left\{ I_\phi^*(\nu) : \nu \in ba(\mathcal{B}(\Delta), \mu),\ A^*\nu = \xi \right\}. \tag{60}$$

Moreover, the sup is attained if $\{\nu \in ba(\mathcal{B}(\Delta), \mu) : A^*\nu = \xi\} \neq \emptyset$.

But, I_ϕ is monotone, therefore $I_\phi^*(\nu) = -\infty$ for all $\nu \notin ba_+(\mathcal{B}(\Delta), \mu)$. Then (60) implies

$$g^*(\xi) = \sup\left\{ I_\phi^*(\nu) : \nu \in ba_+(\mathcal{B}(\Delta), \mu),\ A^*\nu = \xi \right\}. \tag{61}$$

By (59), $I_\phi^*(\nu) = -\infty$ for all $\nu \notin ca(\mathcal{B}(\Delta), \mu)$. Then (61) amounts to $g^*(\xi) = \sup\{I_\phi^*(\nu) : \nu \in ca_+(\mathcal{B}(\Delta), \mu),\ A^*\nu = \xi\} = \sup\{I_\phi^*(\nu) : \nu \in \Gamma(\xi)\}$ and (59) again delivers $g^*(\xi) = A^* I_\phi^*(\xi) = \sup\left\{\int_\Delta \phi^*\left(\frac{d\nu}{d\mu}(\zeta)\right) d\mu(\zeta) : \nu \in \Gamma(\xi)\right\}$. By Corollary 39, $G_\xi(t) = \inf_{k \geq 0}\{kt - g^*(k\xi)\} = \inf_{k\geq 0}\left\{tk - \sup\left\{\int_\Delta \phi^*\left(\frac{d\nu}{d\mu}\right) d\mu : \nu \in \Gamma(k\xi)\right\}\right\}$ for each $(t, \xi) \in \mathbb{R} \times \Delta$, thus, if $\Gamma(\xi) \neq \emptyset$, it follows that

$$\begin{aligned} G_\xi(t) &= \inf_{k\geq 0}\left\{tk - \sup\left\{\int_\Delta \phi^*\left(\frac{d(k\gamma)}{d\mu}(\zeta)\right) d\mu(\zeta) : \gamma \in \Gamma(\xi)\right\}\right\} \\ &= \inf_{\gamma \in \Delta^\sigma(\mu) : \int_\Delta \zeta\, d\gamma(\zeta) = \xi}\ \inf_{k\geq 0}\left\{tk - \int_\Delta \phi^*\left(k\frac{d\gamma}{d\mu}(\zeta)\right) d\mu(\zeta)\right\} \end{aligned}$$

[33] In fact, if ν is countably additive, it is enough to set $u^* = d\nu/d\mu$ to obtain $u^* \in L^1(\mu)$ and $\nu(u) = \int_\Delta u(\zeta)\, d\nu(\zeta) = \int_\Delta u(\zeta)\, u^*(\zeta)\, d\mu(\zeta)$ for all $u \in L^\infty(\mu)$. Conversely, if there exists $u^* \in L^1(\mu)$ such that $\nu(u) = \int_\Delta u(\zeta)\, u^*(\zeta)\, d\mu(\zeta)$ for all $u \in L^\infty(\mu)$, then $\nu(B) = \int_B u^*(\zeta)\, d\mu(\zeta)$ for all $B \in \mathcal{B}(\Delta)$, which implies ν is countably additive and $u^* = d\nu/d\mu$.

else, if $\Gamma(\xi) = \emptyset$, $\sup\left\{\int_\Delta \phi^*\left(\frac{d\nu}{d\mu}\right) d\mu : \nu \in \Gamma(k\xi)\right\} = \sup\left\{\int_\Delta \phi^*\left(\frac{d(k\gamma)}{d\mu}\right) d\mu : \gamma \in \Gamma(\xi)\right\} = -\infty$ if $k > 0$, while $\sup\left\{\int_\Delta \phi^*\left(\frac{d\nu}{d\mu}(\zeta)\right) d\mu(\zeta) : \nu \in \Gamma(k\xi)\right\} = \phi^*(0) = \inf_{k\in\mathbb{R}}\{-\phi(k)\} = -\sup_{k\in\mathbb{R}}\phi(k)$ if $k = 0$, and $G_\xi(t) = \inf_{k\geq 0}\left\{tk - \sup\left\{\int_\Delta \phi^*\left(\frac{d\nu}{d\mu}(\zeta)\right) d\mu(\zeta) : \nu \in \Gamma(k\xi)\right\}\right\} = \sup_{k\in\mathbb{R}}\phi(k)$, which concludes the proof. ■

B.1 Normalized Smooth Preferences Functionals

In this subsection we assume that ϕ is strictly increasing (and concave from $\mathbb{R}$ to $\mathbb{R}$), and consider the normalized version

$$g(x) = \phi^{-1}\left[\int_\Delta \phi(\langle \xi, x\rangle)\, d\mu(\xi)\right] \tag{62}$$

of (56). First observe that $\phi(\mathbb{R})$ is an open half line $(-\infty, a)$, with $a = \sup_{k\in\mathbb{R}}\phi(k)$. Then ϕ^{-1} can be extended to an extended-valued continuous and monotone function from $[-\infty, \infty]$ to $[-\infty, \infty]$ by setting

$$\widehat{\phi^{-1}}(t) = \begin{cases} \infty & \text{if } t \geq a \\ \phi^{-1}(t) & \text{if } a > t > -\infty \\ -\infty & \text{if } t = -\infty \end{cases} \tag{63}$$

this extension is simply denoted ϕ^{-1}. Application of Theorem 55 and Lemma 35 delivers, for all $(t, \xi) \in \mathbb{R}\times\Delta$,

$$G_\xi(t) = \begin{cases} \phi^{-1}\left(\inf\left\{\inf_{k\geq 0}\left[tk - \int_\Delta \phi^*\left(k\frac{d\nu}{d\mu}(\zeta)\right) d\mu(\zeta)\right] : \nu \in \Gamma(\xi)\right\}\right) & \text{if } \Gamma(\xi) \neq \emptyset \\ \phi^{-1}\left(\sup_{k\in\mathbb{R}}\phi(k)\right) = \infty & \text{if } \Gamma(\xi) = \emptyset \end{cases}$$

that is

$$G_\xi(t) = \phi^{-1}\left(\inf\left\{\inf_{k\geq 0}\left[tk - \int_\Delta \phi^*\left(k\frac{d\nu}{d\mu}(\zeta)\right) d\mu(\zeta)\right] : \nu \in \Gamma(\xi)\right\}\right) \tag{64}$$

with the usual convention $\inf\emptyset = \infty$.

Lemma 56 *For a twice differentiable $\phi : \mathbb{R} \to \mathbb{R}$ with $\phi' > 0$ and $\phi'' < 0$, the following facts are equivalent:*

(i) $J_\lambda(\cdot) = \phi(\phi^{-1}(\cdot) + \lambda)$ is concave on $\phi(\mathbb{R})$ for all $\lambda \geq 0$;

(ii) $-\phi'/\phi''$ is weakly decreasing.

In this case ϕ is said to be DARA.

Proposition 57 *If the scalar functions $J_\lambda(r) = \phi[\phi^{-1}(r) + \lambda]$ are concave on $\phi(\mathbb{R})$ for all $\lambda \geq 0$, then (62) is 1-Lipschitz.*

Proof. By the Jensen Inequality, we have $J_\lambda\left(\int_\Delta \phi(\langle\xi,x\rangle)\,d\mu(\xi)\right) \geq \int_\Delta J_\lambda(\phi(\langle\xi,x\rangle))\,d\mu(\xi)$ which implies $g(x)+\lambda \geq g(x+\lambda e)$ for all $\lambda \geq 0$. Proposition 50 delivers 1-Lipschitzianity. ■

Proposition 58 *The functional (62) is translation invariant for all $\mu \in \Delta^\sigma(\mathcal{B}(\Delta))$ if and only if ϕ is CARA.*

Proof. We only prove the "only if," the converse being trivial. If g is translation invariant for all $\mu \in \Delta^\sigma(\mathcal{B}(\Delta))$, then $\phi^{-1}\left(\int_\Delta \phi(\langle\xi,x+\lambda e\rangle)\,d\mu(\xi)\right) = \phi^{-1}\left(\int_\Delta \phi(\langle\xi,x\rangle)\,d\mu(\xi)\right)+\lambda$ for all $x \in X$, $\lambda \in \mathbb{R}$, $\mu \in \Delta^\sigma(\mathcal{B}(\Delta))$. In particular choosing $\xi_1 \neq \xi_2$ in Δ and the probability measure $\mu = (1/2)\,\delta_{\xi_1} + (1/2)\,\delta_{\xi_2}$, we have $\phi^{-1}\left(\frac{\phi(\langle\xi_1,x\rangle+\lambda)+\phi(\langle\xi_2,x\rangle+\lambda)}{2}\right) = \phi^{-1}\left(\frac{\phi(\langle\xi_1,x\rangle)+\phi(\langle\xi_2,x\rangle)}{2}\right)+\lambda$. The linear map $x \mapsto (\langle\xi_1,x\rangle, \langle\xi_2,x\rangle)$ from X into $\mathbb{R}^2$ is onto, because ξ_1 and ξ_2 are linearly independent, therefore $\phi^{-1}\left(\frac{\phi(t+\lambda)+\phi(r+\lambda)}{2}\right) = \phi^{-1}\left(\frac{\phi(t)+\phi(r)}{2}\right)+\lambda$ for all $t, r, \lambda \in \mathbb{R}$. By [17, p. 28] ϕ is CARA. ■

Next we further study the CARA case.

Proposition 59 *The functional $g : X \to \mathbb{R}$ given by*

$$g(x) = -\frac{1}{\theta}\log\int_\Delta e^{-\theta\langle\xi,x\rangle}d\mu(\xi)$$

with $\theta > 0$, is translation invariant and, for every $(t,\xi) \in \mathbb{R}\times\Delta$,

$$g^*(\xi) = -\frac{1}{\theta}\inf\{R(\nu \parallel \mu) : \nu \in \Gamma(\xi)\} \tag{65}$$
$$G_\xi(t) = t + \frac{1}{\theta}\inf\{R(\nu \parallel \mu) : \nu \in \Gamma(\xi)\}.$$

Proof. We first consider the case $\theta = 1$. In view of Theorem 55, let $\phi(t) = -e^{-t}$ and consider the functional $\tilde{g}(x) = \int_\Delta \phi(\langle\xi,x\rangle)\,d\mu(\xi) = \int_\Delta -e^{-\langle\xi,x\rangle}d\mu(\xi)$. Clearly ϕ is concave, increasing, and

$$\phi^*(r) = -\psi^*(r) = \begin{cases} r - r\log r & \text{if } r > 0 \\ 0 & \text{if } r = 0 \\ -\infty & \text{if } r < 0. \end{cases}$$

Simple computation shows that, for all $\nu \in \Delta^\sigma(\mu)$ and $t \in \mathbb{R}$

$$\inf_{k\geq 0}\left[tk - \int_\Delta \phi^*\left(k\frac{d\nu}{d\mu}(\zeta)\right)d\mu(\zeta)\right] = -e^{-t}e^{-R(\nu\|\mu)}.$$

As a consequence, in view of Theorem 55, for all $(t,\xi) \in \mathbb{R}\times\Delta$,

$$\tilde{G}_\xi(t) = \begin{cases} \inf\left\{-e^{-t}e^{-R(\nu\|\mu)} : \nu \in \Gamma(\xi)\right\} & \text{if } \Gamma(\xi) \neq \emptyset \\ \sup_{k\in\mathbb{R}} -e^{-k} & \text{if } \Gamma(\xi) = \emptyset \end{cases}$$
$$= \begin{cases} -e^{-t}\sup\left\{e^{-R(\nu\|\mu)} : \nu \in \Gamma(\xi)\right\} & \text{if } \Gamma(\xi) \neq \emptyset \\ 0 & \text{if } \Gamma(\xi) = \emptyset. \end{cases}$$

Moreover, $g(x) = -\log \int_{\Delta} e^{-\langle \xi, x\rangle} \mu(d\xi) = -\log(-\tilde{g}(x))$, but $r \mapsto -\log(-r)$ is monotone and extended-valued continuous from $[-\infty, 0]$ to $[-\infty, \infty]$. Therefore, if $\Gamma(\xi)$ is not empty, $G_{\xi}(t) = -\log\left(e^{-t} \sup\left\{e^{-R(\nu\|\mu)} : \nu \in \Gamma(\xi)\right\}\right) = t - \sup\left\{\log e^{-R(\nu\|\mu)} : \nu \in \Gamma(\xi)\right\} = t + \inf\{R(\nu \parallel \mu) : \nu \in \Gamma(\xi)\}$, while, if $\Gamma(\xi)$ is empty, then $G_{\xi}(t) = -\log(0) = \infty = t + \inf\{R(\nu \parallel \mu) : \nu \in \Gamma(\xi)\}$. Monotonicity, translation invariance, concavity, and finiteness of g are easily shown. The conjugate of g can then be calculated by (iii) of Corollary 40, thus for all $\xi \in \Delta$, $g^*(\xi) = -G_{\xi}(0) = -\inf\{R(\nu \parallel \mu) : \nu \in \Gamma(\xi)\}$.

Finally, if $\theta \neq 1$, write ${}_{\theta}g$ to emphasize the dependence on θ with ${}_{1}g = g$. Clearly, ${}_{\theta}g(x) = \theta^{-1} g(\theta x)$, therefore $({}_{\theta}g)^* = \frac{1}{\theta} g^*$ and ${}_{\theta}G_{\xi}$ can be calculated by (iii) of Corollary 40. ■

C Proofs

Proof of Lemma 1. It is a direct consequence of Lemma 52. ■

Lemma 60 *If $\succsim$ satisfies A.1 and A.5. Then, $\succsim$ satisfies A.3 if and only if $f, g, h \in \mathcal{F}$, $f \succsim h$, $g \succsim h$, and $\alpha \in (0,1)$ imply $\alpha f + (1-\alpha) g \succsim h$.*

Proof. We prove the "only if" part, the converse being trivial. Suppose A.3 holds. Since $\succsim$ satisfies A.1, to prove the result it is enough to show that $f \succ g$ implies $\alpha f + (1-\alpha) g \succsim g$ for all $\alpha \in (0,1)$. Suppose, *per contra*, that there exist $f \succ g$ and $\bar{\alpha} \in (0,1)$ such that $\bar{\alpha} f + \left(1 - \bar{\alpha}\right) g \prec g$. Then $\bar{\alpha} \in \{\alpha \in [0,1] : g \succsim \alpha f + (1-\alpha) g\} \neq \emptyset$. By A.5, this set is compact. We can therefore set $\beta = \max(\{\alpha \in [0,1] : g \succsim \alpha f + (1-\alpha) g\})$ and $f_{\beta} = \beta f + (1-\beta) g$.

Claim. $f_{\beta} \sim g$.

Proof of the Claim. We have $\beta \in \{\alpha \in [0,1] : g \succsim \alpha f + (1-\alpha) g\}$ and $\beta < 1$. In fact, if $\beta = 1$ then $g \succsim f$, a contradiction. Now suppose $f_{\beta} \nsim g$, that is, $g \succ f_{\beta}$. The set $\{\alpha \in [0,1] : g \succ \alpha f + (1-\alpha) g\}$ is open since it is the complement of the closed set $\{\alpha \in [0,1] : \alpha f + (1-\alpha) g \succsim g\}$. Hence, there is an open neighborhood V in $[0,1]$ containing β and contained in $\{\alpha \in [0,1] : g \succ \alpha f + (1-\alpha) g\}$. Since $\beta < 1$, we can then pick a point $\beta' > \beta$ in V so that $g \succ \beta' f + (1-\beta') g$, which contradicts the maximality of β. We conclude that $f_{\beta} \sim g$ and this completes the proof of the Claim. □

By the Claim, we can apply A.3 to f_{β} and g. Hence, $\lambda f_{\beta} + (1-\lambda) g \succsim g$ for all $\lambda \in (0,1)$, and $0 < \bar{\alpha} < \beta$ implies $\beta^{-1}\bar{\alpha} \in (0,1)$. Thus $g \precsim \frac{\bar{\alpha}}{\beta}(\beta f + (1-\beta) g) + \left(1 - \frac{\bar{\alpha}}{\beta}\right) g = \bar{\alpha} f + \frac{\bar{\alpha}}{\beta} g - \bar{\alpha} g + g - \frac{\bar{\alpha}}{\beta} g = \bar{\alpha} f + \left(1 - \bar{\alpha}\right) g \prec g$, a contradiction. We conclude that $\alpha f + (1-\alpha) g \succsim g$ for all $\alpha \in (0,1)$, as desired. ■

Lemma 61 *A binary relation $\succsim$ on $\mathcal{F}$ satisfies Axiom A.1-A.5 if and only if there exists a nonconstant affine function $u : X \to \mathbb{R}$ and a function $I : B_0(\Sigma, u(X)) \to \mathbb{R}$ normalized, monotone, quasiconcave, and continuous such that*

$$f \succsim g \Longleftrightarrow I(u(f)) \geq I(u(g)). \tag{66}$$

Moreover, u is cardinally unique, and, given u, there is a unique normalized $I : B_0(\Sigma, u(X)) \to \mathbb{R}$ that satisfies (66).

Proof. We only prove the sufficiency of the axioms, the converse being routine. The existence of a nonconstant affine u and a normalized and monotone I satisfying (66) can be derived using the same technique of [38, Lemma 28], where for the existence of u we use Axiom A.4 in place of the stronger Weak Certainty Independence Axiom of [38]. In particular, $B_0(\Sigma, u(X)) = \{u(f) : f \in \mathcal{F}\}$.

By Lemma 60, $\succsim$ is a convex preference, and so I is quasiconcave. Continuity follows from A.5 and Proposition 47.

Finally, cardinal uniqueness of u is a standard result (u is affine and represents $\succsim$ on X). Suppose that, given u, the normalized functionals I_1 and I_2 satisfy (66). For all $\varphi = u(f) \in B_0(\Sigma, u(X))$, let $x_f \in X$ be such that $f \sim x_f$, then $I_1(\varphi) = I_1(u(f)) = I_1(u(x_f)) = u(x_f) = I_2(u(x_f)) = I_2(u(f)) = I_2(\varphi)$, so $I_1 = I_2$. ■

Lemma 62 *Let $\succsim$, I, and u be like in Lemma 61. The following facts are equivalent:*

(i) $\succsim$ satisfies A.7.

(ii) For every $z, z' \in X$, with $z' \prec z$, there are $y' \prec y$ such that, for all $f, g \in \mathcal{F}$

$$\frac{1}{2}f(s) + \frac{1}{2}y' \sim \frac{1}{2}g(s) + \frac{1}{2}y \quad \forall s \in S \Longrightarrow \frac{1}{2}x_f + \frac{1}{2}z' \precsim \frac{1}{2}x_g + \frac{1}{2}z. \tag{67}$$

(iii) I is uniformly continuous.

Proof. Clearly (i) $\Rightarrow$ (ii). Next we show that (ii) $\Rightarrow$ (iii). Let $\varepsilon > 0$ and choose $z, z' \in X$ such that $u(z) - u(z') \leq \varepsilon$ and $0 < u(z) - u(z') < \sup u(X) - \inf u(X)$. Let $y, y' \in X$ be such that (67) is satisfied and set $\delta = u(y) - u(y')$.

Notice that $\delta \in (0, \sup u(X) - \inf u(X))$. Clearly $\delta > 0$, moreover, taking $f = y$ and $g = y'$ we have $\frac{1}{2}f(s) + \frac{1}{2}y' = \frac{1}{2}y + \frac{1}{2}y' \sim \frac{1}{2}y' + \frac{1}{2}y = \frac{1}{2}g(s) + \frac{1}{2}y$ for all $s \in S$, hence $\frac{1}{2}y + \frac{1}{2}z' = \frac{1}{2}x_f + \frac{1}{2}z' \precsim \frac{1}{2}x_g + \frac{1}{2}z = \frac{1}{2}y' + \frac{1}{2}z$, Then $\frac{1}{2}u(y) + \frac{1}{2}u(z') \leq \frac{1}{2}u(y') + \frac{1}{2}u(z)$ and $\delta = u(y) - u(y') \leq u(z) - u(z') < \sup u(X) - \inf u(X)$.

Let $\varphi \in B_0(\Sigma, u(X))$ be such that $\varphi + \delta \in B_0(\Sigma, u(X))$, and $g, f \in \mathcal{F}$ be such that $\varphi = u(g)$ and $\varphi + \delta = u(f)$. Then $u(f(s)) = \varphi(s) + \delta = u(g(s)) + u(y) - u(y')$ for all $s \in S$, $u\left(\frac{1}{2}f(s) + \frac{1}{2}y'\right) = \frac{1}{2}u(f(s)) + \frac{1}{2}u(y') = \frac{1}{2}u(g(s)) + \frac{1}{2}u(y) - \frac{1}{2}u(y') + \frac{1}{2}u(y') = u\left(\frac{1}{2}g(s) + \frac{1}{2}y\right)$ and hence $\frac{1}{2}x_f + \frac{1}{2}z' \precsim \frac{1}{2}x_g + \frac{1}{2}z$, that is $\frac{1}{2}u(x_f) + \frac{1}{2}u(z') \leq \frac{1}{2}u(x_g) + \frac{1}{2}u(z)$ and $I(\varphi + \delta) = I(u(f)) = u(x_f) \leq u(x_g) + (u(z) - u(z')) \leq I(u(g)) + \varepsilon = I(\varphi) + \varepsilon$. Hence, by Proposition 49, I is uniformly continuous.

We conclude by showing that (iii) $\Rightarrow$ (i). Assume I is uniformly continuous. For all $z, z' \in X$, with $z' \prec z$, choose $\delta > 0$ such that $|I(\varphi) - I(\psi)| \leq u(z) - u(z')$ for all $\varphi, \psi \in B_0(\Sigma, u(X))$ such that $\|\varphi - \psi\| \leq \delta$. Take $y' \prec y$ such that $u(y) - u(y') < \delta$. Then for all $f, g \in \mathcal{F}$ such that $\frac{1}{2}f(s) + \frac{1}{2}y' \precsim \frac{1}{2}g(s) + \frac{1}{2}y$ for all $s \in S$, it must be the case that

$$u(f(s)) \leq u(g(s)) + u(y) - u(y') \qquad \forall s \in S. \tag{68}$$

Set $\varphi = u(f)$, $\psi = u(g)$, $\tau = u(y)$, $t = u(y')$, $\delta' = \tau - t \in (0, \delta)$, $\varepsilon = u(z) - u(z')$, $k = \max\{\max \varphi, \max \psi, \tau\} \in u(X)$. Notice that:

- $\varphi \leq (\psi + \delta') \wedge k$. This follows from (68) and the definition of k.
- $(\psi + \delta') \wedge k \in B_0(\Sigma, u(X))$. In fact, $\varphi \leq (\psi + \delta') \wedge k \leq k$.
- $(\psi + \delta') \wedge k = (\psi \wedge (k - \delta')) + \delta'$ and $\psi \wedge (k - \delta') \in B_0(\Sigma, u(X))$. In fact, $k \geq k - \delta' = k - \tau + t \geq \tau - \tau + t = t$.

Therefore $I(u(f)) = I(\varphi) \leq I((\psi + \delta') \wedge k) = I((\psi \wedge (k - \delta')) + \delta')$, but it clearly holds $\|((\psi \wedge (k - \delta')) + \delta') - (\psi \wedge (k - \delta'))\| = \delta' \leq \delta$ and uniform continuity guarantees $I((\psi \wedge (k - \delta')) + \delta') \leq I(\psi \wedge (k - \delta')) + \varepsilon \leq I(\psi) + \varepsilon = I(u(g)) + u(z) - u(z')$, hence $I(u(f)) \leq I(u(g)) + u(z) - u(z')$ and $u(x_f) - u(x_g) \leq u(z) - u(z')$ which amounts to $\frac{1}{2}x_f + \frac{1}{2}z' \precsim \frac{1}{2}x_g + \frac{1}{2}z$, as wanted. ■

Lemma 63 *Let $\succsim$ be a binary relation on X represented by an affine function $u : X \to \mathbb{R}$. $u(X) = \mathbb{R}$ if and only if $\succsim$ satisfies A.6.*

Proof of Theorem 3. Suppose (i) holds, i.e., $\succsim$ satisfies Axioms A.1-A.5. By Lemma 61, there exists a nonconstant affine function $u : X \to \mathbb{R}$ and a function $I : B_0(\Sigma, u(X)) \to \mathbb{R}$ normalized, monotone, quasiconcave, and continuous such that $f \succsim g \iff I(u(f)) \geq I(u(g))$. By Corollary 38, $I(\varphi) = \inf_{p \in \Delta} G_p(\langle p, \varphi \rangle)$ for all $\varphi \in B_0(\Sigma, u(X))$, i.e. $I(u(f)) = \inf_{p \in \Delta} G_p\left(\int u(f)\, dp\right)$ for all $f \in \mathcal{F}$, where $G_p(t) = \sup\{I(\varphi) : \varphi \in B_0(\Sigma, u(X)), \langle p, \varphi \rangle \leq t\}$ for all $(t, p) \in u(X) \times \Delta$.[34]

[34] Indeed $G_p(t)$ is defined for all $(t, p) \in \mathbb{R} \times \Delta$, but notice that $\langle p, \varphi \rangle \in u(X)$ for all $\varphi \in B_0(\Sigma, u(X))$.

Lemma 31 implies that the map $(t,p)\mapsto G_p(t)$ is quasiconvex on $u(X)\times\Delta$. Monotonicity of $G_p(\cdot)$ is obvious. Moreover, for all $t\in u(X)$, $t=I(t)=\inf_{p\in\Delta}G_p(\langle p,t\rangle)=\inf_{p\in\Delta}G_p(t)$. Therefore, $G^\star:u(X)\times\Delta\to(-\infty,\infty]$ defined by $G^\star(t,p)=G_p(t)$ is well defined (the above equation rules out the value $-\infty$), belongs to $\mathcal{G}(u(X)\times\Delta)$, is linearly continuous because of continuity of I, and (9) holds. This proves (ii).

Conversely, suppose (ii) holds. Since $G\in\mathcal{G}(u(X)\times\Delta)$, then, by Lemma 51,

$$I(\varphi)=\inf_{p\in\Delta}G(\langle p,\varphi\rangle,p)\qquad\forall\varphi\in B_0(\Sigma,u(X))\tag{69}$$

is finite, (evenly) quasiconcave, monotone, normalized. Linear continuity of G implies continuity of I, and (9) amounts to

$$f\succsim g\Longleftrightarrow I(u(f))\geq I(u(g)).\tag{70}$$

Lemma 61 guarantees that $\succsim$ satisfies A.1-A.5, i.e., (i) holds.

Assume (i), or (ii), holds and $v:X\to\mathbb{R}$ is nonconstant affine, $H\in\mathcal{G}(v(X)\times\Delta)$, for all f and g in $\mathcal{F}$,

$$f\succsim g\Longleftrightarrow\inf_{p\in\Delta}H\left(\int v(f)\,dp,p\right)\geq\inf_{p\in\Delta}H\left(\int v(g)\,dp,p\right).\tag{71}$$

Notice that we are not requiring that H be linearly continuous. Define

$$J(\varphi)=\inf_{p\in\Delta}H(\langle p,\varphi\rangle,p)\qquad\forall\varphi\in B_0(\Sigma,v(X)).\tag{72}$$

Since $H\in\mathcal{G}(v(X)\times\Delta)$, then, by Lemma 51, J is finite, (evenly) quasiconcave, monotone, normalized,

$$H(t,p)\geq\sup\{J(\varphi):\varphi\in B_0(\Sigma,v(X))\text{ and }\langle p,\varphi\rangle\leq t\}\qquad\forall(t,p)\in v(X)\times\Delta\tag{73}$$

and (71) amounts to

$$f\succsim g\Longleftrightarrow J(v(f))\geq J(v(g)).\tag{74}$$

Since J is normalized, by (74), v represents $\succsim$ on X, then it is cardinally equivalent to u. Assume $v=u$, then (70), (74), and Lemma 61 guarantee that $J=I$ (in particular H is linearly continuous too). By (73), for all $(t,p)\in u(X)\times\Delta$, $H(t,p)\geq\sup\{I(\varphi):\varphi\in B_0(\Sigma,u(X))\text{ and }\langle p,\varphi\rangle\leq t\}=G_p(t)$. Since I is finite, normalized, monotone, quasiconcave, and continuous, we can proceed verbatim like in the proof that (i) implies (ii) (starting from "By Corollary 38...") to show that $G^\star:u(X)\times\Delta\to(-\infty,\infty]$ defined by $G^\star(t,p)=G_p(t)$ is well defined, belongs to $\mathcal{G}(u(X)\times\Delta)$, is linearly continuous, and

$$f\succsim g\Longleftrightarrow\inf_{p\in\Delta}G^\star\left(\int u(f)\,dp,p\right)\geq\inf_{p\in\Delta}G^\star\left(\int u(g)\,dp,p\right).\tag{75}$$

Thus $(u, G^\star)$ represents $\succsim$ in the sense of (ii) and $G^\star$ is the minimal element of $\mathcal{G}(u(X) \times \Delta)$ with this property. Moreover, for all $(t,p) \in u(X) \times \Delta$, $\sup_{f \in \mathcal{F}} \left\{ u(x_f) : \int u(f)\, dp \leq t \right\} = \sup_{f \in \mathcal{F}} \left\{ I(u(f)) : \int u(f)\, dp \leq t \right\} = \sup \{ I(\varphi) : \varphi \in B_0(\Sigma, u(X)) \text{ and } \langle p, \varphi \rangle \leq t \} = G_p(t) = G^\star(t,p)$.

Finally, it is easy to check that $\succsim$ has no worst consequence if and only if $\inf u(X) \notin u(X)$. In this case, $B_0(\Sigma, u(X))$ is lower open. By Lemma 32, the map $(t,p) \mapsto G_p(t)$ is lower semicontinuous on $u(X) \times \Delta$, thus $p \mapsto G_p(\langle p, \varphi \rangle)$ is lower semicontinuous on Δ, and the infima in (75) are attained. ■

Proof of Proposition 4. Let (u, G) be an uncertainty averse representation of a preference $\succsim$. If $(\bar{u}, \bar{G})$ is another uncertainty averse representation of $\succsim$, then by standard uniqueness results, there exist $\alpha > 0$ and $\beta \in \mathbb{R}$ such that $\bar{u} = \alpha u + \beta$. By (10), for all $(t,p) \in \bar{u}(X) \times \Delta$,

$$\begin{aligned} \bar{G}(t,p) &= \sup_{f \in \mathcal{F}} \left\{ \bar{u}(x_f) : \int \bar{u}(f)\, dp \leq t \right\} = \sup_{f \in \mathcal{F}} \left\{ \alpha u(x_f) + \beta : \alpha \int u(f)\, dp + \beta \leq t \right\} \\ &= \alpha \sup_{f \in \mathcal{F}} \left\{ u(x_f) : \int u(f)\, dp \leq \frac{t-\beta}{\alpha} \right\} + \beta = \alpha G\left(\frac{t-\beta}{\alpha}, p \right) + \beta \end{aligned}$$

as desired. The converse is a simple verification. ■

Proof of Theorem 5. Let $\succsim$ be uncertainty averse and satisfy Axioms A.4-A.6. Assume $u : X \to \mathbb{R}$ is affine, $G \in \mathcal{G}(u(X) \times \Delta)$ is lower semicontinuous, and, for all f and g in $\mathcal{F}$, (9) holds. Then, by A.6, $u(X) = \mathbb{R}$ (see Lemma 63). Set $I(\varphi) = \inf_{p \in \Delta} G(\langle p, \varphi \rangle, p)$ for all $\varphi \in B_0(\Sigma)$. Since $G \in \mathcal{G}(\mathbb{R} \times \Delta)$, then, by Lemma 51, I is finite, (evenly) quasiconcave, monotone, normalized, and (9) amounts to $f \succsim g \iff I(u(f)) \geq I(u(g))$ for all f and g in $\mathcal{F}$. Since $G \in \mathcal{G}(\mathbb{R} \times \Delta)$ is lower semicontinuous, then it satisfies the assumptions of Lemma 42, and $G(t,p) = \sup_{\varphi \in B_0(\Sigma) : \langle p, \varphi \rangle \leq t} I(\varphi) = \sup_{f \in \mathcal{F} : \langle p, u(f) \rangle \leq t} I(u(f))$ for all $(t,p) \in \mathbb{R} \times \Delta$. But, since I is normalized and $I(u(\cdot))$ represents $\succsim$, then $I(u(f)) = u(x_f)$ for all $f \in \mathcal{F}$ (notice that the existence of x_f is guaranteed by A.1, A.2, and A.5), therefore $G(t,p) = \sup_{f \in \mathcal{F} : \langle p, u(f) \rangle \leq t} I(u(f)) = \sup_{f \in \mathcal{F}} \left\{ u(x_f) : \int u(f)\, dp \leq t \right\}$ for all $(t,p) \in \mathbb{R} \times \Delta$. This proves that that (10) holds, and $G = G^\star$. ■

Proof of Proposition 6. By standard results ([25, Corollary B.3]), (i) implies that $u_1 \approx u_2$. Wlog, $u_1 = u_2 = u$. By (12), for all $f \in \mathcal{F}$ and $x \in X$, $f \sim_1 x$ implies $f \succsim_2 x$, and so $x_f^2 \sim_2 f \succsim_2 x_f^1$ (where $f \sim_i x_f^i \in X$, for $i = 1, 2$). Hence, $u(x_f^2) \geq u(x_f^1)$ for all $f \in \mathcal{F}$. By (10), for all $(t,p) \in u(X) \times \Delta$, $G_1(t,p) = \sup_{f \in \mathcal{F}} \left\{ u(x_f^1) : \int u(f)\, dp \leq t \right\} \leq \sup_{f \in \mathcal{F}} \left\{ u(x_f^2) : \int u(f)\, dp \leq t \right\} = G_2(t,p)$, and so $G_1 \leq G_2$.

Conversely, assume wlog $u_1 = u_2 = u$. Then, for all $f \in \mathcal{F}$ and $x \in X$, $f \succsim_1 x$ implies $\inf_{p \in \Delta} G_1\left(\int u(f)\, dp, p \right) \geq \inf_{p \in \Delta} G_1\left(\int u(x)\, dp, p \right) = u(x)$, but $G_1 \leq G_2$ implies $\inf_{p\in\Delta} G_2\left(\int u(f)\,dp, p\right) \geq \inf_{p\in\Delta} G_1\left(\int u(f)\,dp, p\right) \geq u(x) = \inf_{p\in\Delta} G_2\left(\int u(x)\,dp, p\right)$, which delivers $f \succsim_2 x$. ■

Proof of Theorem 7. Suppose $\succsim$ satisfies Axioms A.1-A.6. By Lemma 61, there exists a nonconstant affine function $u : X \to \mathbb{R}$ and a function $I : B_0(\Sigma, u(X)) \to \mathbb{R}$ normalized, monotone, quasiconcave, and continuous such that $f \succsim g \Longleftrightarrow I(u(f)) \geq I(u(g))$. Moreover, by Lemma 63, $u(X) = \mathbb{R}$. Then $B_0(\Sigma, u(X)) = B_0(\Sigma)$. Set $G_p(t) = \sup\{I(\varphi) : \varphi \in B_0(\Sigma) \text{ and } \langle p, \varphi\rangle \leq t\}$ for all $(t,p) \in \mathbb{R}\times\Delta$. Theorem 36 guarantees that $I(\varphi) = \min_{p\in\Delta} G_p(\langle p, \varphi\rangle)$ for all $\varphi \in B_0(\Sigma)$. In particular, $I(u(f)) = \min_{p\in\Delta} G_p\left(\int u(f)\,dp\right)$ for all $f \in \mathcal{F}$, and (13) holds. Lemmas 31 and 32 guarantee that the map $(t,p) \mapsto G_p(t)$ is quasiconvex and lower semicontinuous on $\mathbb{R}\times\Delta$. Monotonicity of $G_p(\cdot)$ is obvious. Moreover, for all $t \in \mathbb{R}$, $t = I(t) = \min_{p\in\Delta} G_p(\langle p, t\rangle) = \min_{p\in\Delta} G_p(t)$. Therefore, $G^\star : \mathbb{R}\times\Delta \to (-\infty, \infty]$ defined by $G^\star(t,p) = G_p(t)$ is well defined, lower semicontinuous, and it belongs to $\mathcal{G}(\mathbb{R}\times\Delta)$. Since I is continuous, $G^\star$ is linearly continuous. By Theorem 5, $(u, G^\star)$ is an uncertainty averse representation of $\succsim$.

If $\succsim$ also satisfies Axiom A.7, that is (i) holds, by Lemma 62, I is uniformly continuous. Then, by Theorem 41, $\operatorname{dom}(G_p) \in \{\emptyset, \mathbb{R}\}$ for all $p \in \Delta$, and $\{G_p\}_{p\in\Delta:\operatorname{dom}(G_p)=\mathbb{R}}$ are uniformly equicontinuous, implying $G \in \mathcal{E}(\mathbb{R}\times\Delta)$ and hence (ii).

Conversely, suppose $G \in \mathcal{G}(\mathbb{R}\times\Delta)$ is lower semicontinuous and linearly continuous and u is affine and onto. Since $G \in \mathcal{G}(\mathbb{R}\times\Delta)$, then, by Lemma 51,

$$I(\varphi) = \inf_{p\in\Delta} G(\langle p, \varphi\rangle, p) \qquad \forall \varphi \in B_0(\Sigma) \tag{76}$$

is finite, (evenly) quasiconcave, monotone, normalized, and (13) amounts to

$$f \succsim g \Longleftrightarrow I(u(f)) \geq I(u(g)). \tag{77}$$

Since G is linearly continuous, I is continuous, thus, by Lemma 61, $\succsim$ satisfies A.1-A.5. Since u is affine, $u(X) = \mathbb{R}$, and u represents $\succsim$ on X, then Lemma 63 guarantees that $\succsim$ satisfies A.6. By Theorem 5, (u, G) is an uncertainty averse representation of $\succsim$.

If $G \in \mathcal{E}(\mathbb{R}\times\Delta)$, that is (ii) holds, then G satisfies the previous properties. Hence, $\succsim$ satisfies A.1-A.6. Further, by Lemma 53, $G \in \mathcal{E}(\mathbb{R}\times\Delta)$ implies that I is uniformly continuous, thus, by Lemma 62, $\succsim$ satisfies A.7 too. This proves (i).

From this point until the end of the proof we (only) assume $\succsim$ satisfies Axioms A.1-A.6 and denote: by (u, G) an uncertainty averse representation, and by I the functional defined in (76).

By Theorem 3, u is cardinally unique, by definition of uncertainty averse representation, for all $(t,p) \in \mathbb{R}\times\Delta$, $G(t,p) = \sup_{f\in\mathcal{F}}\left\{u(x_f) : \int u(f)\,dp \leq t\right\} = \sup\{I(\varphi) : \varphi \in B_0(\Sigma)$ and

$\langle p,\varphi\rangle \leq t\} = \sup\{I(\varphi) : \varphi \in B_0(\Sigma) \text{ and } \langle p,\varphi\rangle = t\} = \sup_{f\in\mathcal{F}}\left\{I(u(f)) : \int u(f)\,dp = t\right\} = \sup_{f\in\mathcal{F}}\left\{u(x_f) : \int u(f)\,dp = t\right\}$, where the central equality descends from Lemma 34. This proves that, given u, G is unique and that (14) holds.

Finally, assume Σ is a σ-algebra. If $\succsim$ satisfies Axiom A.8, assume $\varphi,\psi \in B_0(\Sigma)$, $k \in \mathbb{R}$, $\Sigma \ni E_n \downarrow \emptyset$, and $I(\psi) > I(\varphi)$. Choose $f, g \in \mathcal{F}$ and $x \in X$ such that $\varphi = u(g)$, $\psi = u(f)$, and $k = u(x)$, then $f \succ g$ and there exists $n \in \mathbb{N}$ such that $xE_nf \succ g$, that is $I\left(k1_{E_n} + \psi 1_{E_n^c}\right) = I\left(u(x)1_{E_n} + u(f)1_{E_n^c}\right) = I(u(xE_nf)) > I(u(g)) = I(\varphi)$. By Theorem 54, there is $q \in \Delta^\sigma$ such that $G(\cdot,p) \equiv \infty$ for all $p \notin \Delta^\sigma(q)$, thus the minima in (13) are attained in $\Delta^\sigma(q)$. Conversely, if there is $q \in \Delta^\sigma$ such that $G(\cdot,p) \equiv \infty$ for all $p \notin \Delta^\sigma(q)$, by Theorem 54, for all $\varphi,\psi \in B_0(\Sigma)$, $k \in \mathbb{R}$, $\Sigma \ni E_n \downarrow \emptyset$, $I(\psi) > I(\varphi)$ implies that there exists $n \in \mathbb{N}$ such that $I\left(k1_{E_n} + \psi 1_{E_n^c}\right) > I(\varphi)$. Let $f \succ g$ in $\mathcal{F}$, $x \in X$, and $\Sigma \ni E_n \downarrow \emptyset$, then $\varphi = u(g), \psi = u(f) \in B_0(\Sigma)$, $k = u(x) \in \mathbb{R}$, and $I(\psi) = I(u(f)) > I(u(g)) = I(\varphi)$. Then there exists $n \in \mathbb{N}$ such that $I\left(k1_{E_n} + \psi 1_{E_n^c}\right) > I(\varphi)$, but $I\left(k1_{E_n} + \psi 1_{E_n^c}\right) = I\left(u(x)1_{E_n} + u(f)1_{E_n^c}\right) = I(u(xE_nf))$ and $I(\varphi) = I(u(g))$, thus $I(u(xE_nf)) > I(u(g))$ and $xE_nf \succ g$. In conclusion, A.8 holds. ■

Proof of Theorem 10. First notice that $\mathrm{dom}_\Delta G$ is convex. If $q \notin C^*$, there exist $\varphi_0 \in B_0(\Sigma)$, $r \in \mathbb{R}$ and $\varepsilon > 0$ such that $\int \varphi_0 dq < r < r + \varepsilon < \int \varphi_0 dp$ for all $p \in C^*$. Replacing φ_0 with $\varphi_0 - r$ allows to assume $r = 0$. In this case, for each $x \in X$ and $t \in \mathbb{R}$ there is $n \in \mathbb{N}$ such that $\int n\varphi_0 dq < t$ and $\int n\varphi_0 dp > u(x)$ for all $p \in C^*$. Taking $g \in \mathcal{F}$ such that $u(g) = n\varphi_0$, $\int u(g)\,dq < t$ and $g \succsim^* x$. Therefore $g \in \left\{f \in \mathcal{F} : \int u(f)\,dq \leq t\right\}$ and $u(x_g) \geq u(x)$, since $g \succsim^* x$ implies $g \succsim x$.

Summing up: If $q \notin C^*$, for each $x \in X$ and $t \in \mathbb{R}$ there is $g \in \mathcal{F}$ such that $g \in \left\{f \in \mathcal{F} : \int u(f)\,dq \leq t\right\}$ and $u(x_g) \geq u(x)$, that is $G(t,q) = \infty$. Then $(C^*)^c \subseteq (\mathrm{dom}_\Delta G)^c$, that is $\mathrm{dom}_\Delta G \subseteq C^*$, and $\mathrm{cl}(\mathrm{dom}_\Delta G) \subseteq C^*$, since C^* is closed.

Conversely, let $D = \mathrm{dom}_\Delta G$, and $\varphi,\psi \in B_0(\Sigma)$. Then choosing $f, g \in \mathcal{F}$ such that $\varphi = u(f)$ and $\psi = u(g)$,

$$
\begin{aligned}
&\int \varphi dq \geq \int \psi dq \quad \forall q \in D \\
&\Rightarrow \int u(\alpha f + (1-\alpha)h)\,dq \geq \int u(\alpha g + (1-\alpha)h)\,dq \quad \forall q \in D, \alpha \in [0,1], h \in \mathcal{F} \\
&\Rightarrow \min_{q\in\Delta} G\left(\int u(\alpha f + (1-\alpha)h)\,dq, q\right) \geq \min_{q\in\Delta} G\left(\int u(\alpha g + (1-\alpha)h)\,dq, q\right) \ \forall \alpha \in (0,1], h \in \mathcal{F} \\
&\Rightarrow \alpha f + (1-\alpha)h \succsim \alpha g + (1-\alpha)h \quad \forall \alpha \in (0,1], h \in \mathcal{F} \\
&\Rightarrow \int \varphi dp \geq \int \psi dp \quad \forall p \in C^* .
\end{aligned}
$$

[25, Prop. A.1] implies $C^* \subseteq \mathrm{cl}(\mathrm{co}(\mathrm{dom}_\Delta G)) = \mathrm{cl}(\mathrm{dom}_\Delta G)$. ■

Proof of Proposition 11. (i). For the proof of point (i), we will not need A.6. Define $I : B_0(\Sigma, u(X)) \to \mathbb{R}$ by $I(\varphi) = \inf_{p\in\Delta} G\left(\int \varphi dp, p\right)$ for all $\varphi \in B_0(\Sigma, u(X))$. I is monotone, quasiconcave, continuous on $B_0(\Sigma, u(X))$ and can be naturally extended to $B_0(\Sigma)$ by putting $I(\psi) = -\infty$ if $\psi \notin B_0(\Sigma, u(X))$.

Notice that $G(t,p) = \sup_{f\in\mathcal{F}} \left\{I(u(f)) : \int u(f)\, dp \le t\right\} = \sup_{\varphi\in B_0(\Sigma)} \left\{I(\varphi) : \int \varphi dp \le t\right\}$. Moreover, dom $I = B_0(\Sigma, u(X))$, and, for all $\varphi_0 \in$ dom I, the normalized Greenberg-Pierskalla superdifferential of I at φ_0 is $\partial^\star I(\varphi_0) = \{\mu \in ba(\Sigma) : [I > I(\varphi_0)] \subseteq [\mu > \mu(\varphi_0)]\}$.

Now let $f_0 \in \mathcal{F}$ and set $\varphi_0 = u(f_0)$. Simple algebra delivers

$$\begin{aligned}\partial^\star I(u(f_0)) \cap \Delta &= \left\{p \in \Delta : f \succ f_0 \text{ implies } \int u(f)\, dp > \int u(f_0)\, dp\right\} \\ &= \left\{p \in \Delta : \int u(f)\, dp \le \int u(f_0)\, dp \text{ implies } f_0 \succsim f\right\} = \pi_u(f_0).\end{aligned}$$

Next we show that $\arg\inf_{p\in\Delta} G\left(\int u(f_0)\, dp, p\right) = \partial^\star I(u(f_0)) \cap \Delta$.
$\subseteq$) If $p_0 \in \arg\inf_{p\in\Delta} G\left(\int u(f_0)\, dp, p\right)$, then $p_0 \in \Delta$ and we have $I(\varphi_0) = G\left(\int \varphi_0 dp_0, p_0\right) = \sup_{\varphi\in B_0(\Sigma)} \left\{I(\varphi) : \int \varphi dp_0 \le \int \varphi_0 dp_0\right\}$, that is

$$\begin{gathered} I(\varphi_0) \ge I(\varphi) \text{ for all } \varphi \in B_0(\Sigma) \text{ s.t. } \int \varphi dp_0 \le \int \varphi_0 dp_0 \\ \left[\varphi \in B_0(\Sigma) : \int \varphi dp_0 \le \int \varphi_0 dp_0\right] \subseteq [\varphi \in B_0(\Sigma) : I(\varphi) \le I(\varphi_0)] \\ [\varphi \in B_0(\Sigma) : I(\varphi) > I(\varphi_0)] \subseteq \left[\varphi \in B_0(\Sigma) : \int \varphi dp_0 > \int \varphi_0 dp_0\right]\end{gathered}$$

thus $p_0 \in \partial^\star I(u(f_0))$.
$\supseteq$) If $p_0 \in \partial^\star I(u(f_0)) \cap \Delta$, then $p_0 \in \Delta$ and $[I \le I(\varphi_0)] \supseteq \left[p_0 \le \int \varphi_0 dp_0\right]$ thus

$$\inf_{p\in\Delta} G\left(\int \varphi_0 dp, p\right) = I(\varphi_0) = \sup_{\varphi\in B_0(\Sigma)} \left\{I(\varphi) : \int \varphi dp_0 \le \int \varphi_0 dp_0\right\} = G\left(\int \varphi_0 dp_0, p_0\right)$$

as wanted.

(ii). Consider the binary relation $\succsim^{**}$ defined by

$$f \succsim^{**} g \text{ if and only if } \int u(f)\, dp \ge \int u(g)\, dp \quad \forall p \in \bigcup_{f'\in\mathcal{F}} \pi_u(f'). \tag{78}$$

If $f \succsim^{**} g$ then $\int u(\alpha f + (1-\alpha) h)\, dp \ge \int u(\alpha g + (1-\alpha) h)\, dp$ for all $\alpha \in (0,1]$, $h \in \mathcal{F}$, and $p \in \bigcup_{f'\in\mathcal{F}} \pi_u(f')$. Taking, for each $\alpha \in (0,1]$ and for each $h \in \mathcal{F}$, $\bar{p} = \bar{p}_{\alpha,h} \in \pi_u(\alpha f + (1-\alpha) h)$ it follows that

$$\begin{aligned} I(u(\alpha f + (1-\alpha) h)) &= G\left(\int u(\alpha f + (1-\alpha) h)\, d\bar{p}, \bar{p}\right) \ge G\left(\int u(\alpha g + (1-\alpha) h)\, d\bar{p}, \bar{p}\right) \\ &\ge I(u(\alpha g + (1-\alpha) h)).\end{aligned}$$

Thus, $f \succsim^{**} g$ implies that $f \succsim^{*} g$. By [25, Proposition A.1.], it then follows that $C^* \subseteq \mathrm{cl}\left(\mathrm{co}\left(\bigcup_{f' \in \mathcal{F}} \pi_u(f')\right)\right)$. Conversely, if $p \in \bigcup_{f \in \mathcal{F}} \pi_u(f)$, then there is $f_p \in \mathcal{F}$ such that $\mathbb{R} \ni I(u(f_p)) = G\left(\int u(f_p)\,dp, p\right)$, thus $p \in \mathrm{dom}_\Delta G$. Hence, $\mathrm{cl}\left(\mathrm{co}\left(\bigcup_{f \in \mathcal{F}} \pi_u(f)\right)\right) \subseteq \mathrm{cl}(\mathrm{dom}_\Delta G) = C^*$. ■

Proof of Proposition 12. If (u, c) is a variational representation of $\succsim$, it is routine to check that (u, G) is a representation in the sense of (9). Moreover, since $u(X) = \mathbb{R}$, then Theorem 5 guarantees that (u, G) is an uncertainty averse representation.

Conversely, if (u, G) is an uncertainty averse representation of $\succsim$, and there exist $\gamma : \mathbb{R} \to \mathbb{R}$ and $c : \Delta \to [0, \infty]$ with $\inf_{p \in \Delta} c(p) = 0$, such that $G(t, p) = \gamma(t) + c(p)$ for all $(t, p) \in \mathbb{R} \times \Delta$, then for all $t \in \mathbb{R}$, $t = \inf_{p \in \Delta} [\gamma(t) + c(p)] = \gamma(t) + \inf_{p \in \Delta} c(p) = \gamma(t)$. Hence, γ is the identity. Moreover, if $p_\alpha \to p$ in Δ, then $(0, p_\alpha) \to (0, p)$ in $\mathbb{R} \times \Delta$ and lower semicontinuity of G delivers $\liminf_\alpha c(p_\alpha) = \liminf_\alpha G(0, p_\alpha) \geq G(0, p) = c(p)$, thus c is lower semicontinuous.

Finally the quasiconvexity of G implies that c is convex. In fact, let p_1 and p_2 in $\mathrm{dom}(c)$ and $\alpha \in (0, 1)$. Pick $t_2, t_1 \in \mathbb{R}$ so that $c(p_1) - c(p_2) = t_2 - t_1$, namely, $t_1 + c(p_1) = t_2 + c(p_2)$. As $G : (t, p) \to t + c(p)$ is quasiconvex, then $\alpha t_1 + (1 - \alpha) t_2 + c(\alpha p_1 + (1 - \alpha) p_2) \leq \max\{t_1 + c(p_1), t_2 + c(p_2)\} = t_2 + c(p_2)$, hence $c(\alpha p_1 + (1 - \alpha) p_2) \leq c(p_2) + t_2 - \alpha t_1 - (1 - \alpha) t_2 = c(p_2) + t_2 - \alpha t_1 - t_2 + \alpha t_2 = c(p_2) + \alpha(t_2 - t_1) = c(p_2) + \alpha(c(p_1) - c(p_2)) = \alpha c(p_1) + (1 - \alpha) c(p_2)$, as wanted. ■

Proof of Theorem 19. Assume (21) holds. Set

$$J(\varphi) = \int_\Delta \phi(\langle p, \varphi \rangle)\,d\mu(p) \in \phi(\mathbb{R}) \qquad \forall \varphi \in B_0(\Sigma). \tag{79}$$

By Theorem 55, J is finite, concave, continuous and monotone on X. Therefore the functional

$$I = \phi^{-1} \circ J \tag{80}$$

is well defined, quasiconcave, continuous, monotone, and normalized. Moreover, by (21), $f \succsim g$ if and only if $I(u(f)) \geq I(u(g))$. Thus $\succsim$ satisfies Axioms A.1-A.5 and its uncertainty averse representation (u, G) corresponding to u satisfies, for all $(t, p) \in \mathbb{R} \times \Delta$, $G(t, p) = \sup_{f \in \mathcal{F}} \left\{u(x_f) : \int u(f)\,dp \leq t\right\} = \sup_{f \in \mathcal{F} : \langle p, u(f) \rangle \leq t} I(u(f)) = \sup_{\varphi \in B_0(\Sigma) : \langle p, \varphi \rangle \leq t} I(\varphi)$, by (64), $\sup_{\varphi \in B_0(\Sigma) : \langle p, \varphi \rangle \leq t} I(\varphi) = \phi^{-1}\left(\inf\left\{\inf_{k \geq 0}\left[tk - \int_\Delta \phi^*\left(k \frac{d\nu}{d\mu}(q)\right) d\mu(q)\right] : \nu \in \Gamma(p)\right\}\right)$ and $\Gamma(p) =$

$\left\{\nu \in \Delta^\sigma\left(\mathcal{B}\left(\Delta\right), \mu\right) : p = \int_\Delta q d\nu\left(q\right)\right\}$; by definition of $I_t\left(\cdot \parallel \mu\right)$,

$$\begin{aligned}
&\phi^{-1}\left(\inf\left\{\inf_{k\geq 0}\left[tk - \int_\Delta \phi^*\left(k\frac{d\nu}{d\mu}\left(q\right)\right)d\mu\left(q\right)\right] : \nu \in \Gamma\left(p\right)\right\}\right) \\
&= \inf_{\nu\in\Gamma(p)} \phi^{-1}\left(\inf_{k\geq 0}\left[tk - \int_\Delta \phi^*\left(k\frac{d\nu}{d\mu}\left(q\right)\right)d\mu\left(q\right)\right]\right) = \inf_{\nu\in\Gamma(p)}\left\{I_t\left(\nu \parallel \mu\right) + t\right\} \\
&= t + \inf_{\nu\in\Gamma(p)} I_t\left(\nu \parallel \mu\right)
\end{aligned}$$

and the infimum is attained since $\Gamma\left(p\right)$ is weakly closed and $I_t\left(\cdot \parallel \mu\right)$ has weakly compact sublevel sets. That is $G\left(t,p\right) = t + \min_{\nu\in\Gamma(p)} I_t\left(\nu \parallel \mu\right)$ for all $\left(t,p\right) \in \mathbb{R}\times\Delta$.

Conversely, assume (u, G) is an uncertainty averse representation of $\succsim$, where

$$G\left(t,p\right) = t + \min_{\nu\in\Gamma(p)} I_t\left(\nu \parallel \mu\right) \tag{81}$$

for all $\left(t,p\right) \in \mathbb{R}\times\Delta$, with $\Gamma\left(p\right) = \left\{\nu \in \Delta^\sigma\left(\mathcal{B}\left(\Delta\right), \mu\right) : p = \int_\Delta q d\nu\left(q\right)\right\}$, under the convention $G\left(\cdot, p\right) \equiv \infty$ when $\Gamma\left(p\right) = \emptyset$. Then, for all $\left(t,p\right) \in \mathbb{R}\times\Delta$,

$$G\left(t,p\right) = \phi^{-1}\left(\inf_{\nu\in\Gamma(p)}\left\{\inf_{k\geq 0}\left[kt - \int_\Delta \phi^*\left(k\frac{d\nu}{d\mu}\right)d\mu\right]\right\}\right)$$

and defining J and I like in (79) and (80), it descends from (64) that, for all $\left(t,p\right) \in \mathbb{R}\times\Delta$, $G\left(t,p\right) = \sup_{\varphi\in B_0(\Sigma):\langle p,\varphi\rangle\leq t} I\left(\varphi\right)$. Since I is finite, quasiconcave, continuous, monotone, and normalized, by Theorem 36,

$$I\left(\varphi\right) = \inf_{p\in\Delta} G\left(\left\langle p,\varphi\right\rangle, p\right) = \min_{p\in\Delta} G\left(\left\langle p,\varphi\right\rangle, p\right) \qquad \forall\varphi \in B_0\left(\Sigma\right). \tag{82}$$

Since (u, G) is an uncertainty averse representation of $\succsim$, then, for all f and g in $\mathcal{F}$,

$$\begin{aligned}
f \succsim g &\iff \inf_{p\in\Delta} G\left(\int u\left(f\right)dp, p\right) \geq \inf_{p\in\Delta} G\left(\int u\left(g\right)dp, p\right) \\
&\iff \min_{p\in\Delta} G\left(\int u\left(f\right)dp, p\right) \geq \min_{p\in\Delta} G\left(\int u\left(g\right)dp, p\right) \\
&\iff I\left(u\left(f\right)\right) \geq I\left(u\left(g\right)\right) \iff \phi\left(I\left(u\left(f\right)\right)\right) \geq \phi\left(I\left(u\left(g\right)\right)\right) \\
&\iff \int_\Delta \phi\left(\int_S u\left(f\left(s\right)\right)dp\left(s\right)\right)d\mu\left(p\right) \geq \int_\Delta \phi\left(\int_S u\left(g\left(s\right)\right)dp\left(s\right)\right)d\mu\left(p\right)
\end{aligned}$$

as wanted. Finally notice that (81) and (82) imply (27). ■

Proof of Proposition 20. First observe that, for each $p \in \Delta$,

$$\Gamma\left(p\right) = \left\{\nu \in \Delta^\sigma\left(\mathcal{B}\left(\Delta\right), \mu\right) : \int_S \varphi dp = \int_\Delta\left(\int_S \varphi dq\right)d\nu\left(q\right) \quad \forall\varphi \in B_0\left(\Sigma\right)\right\}.$$

Moreover, $\{\Gamma(p)\}_{p\in\Delta:\Gamma(p)\neq\emptyset}$ coincides with the family of equivalence classes of the equivalence relation R on $\Delta^{\sigma}(\mathcal{B}(\Delta),\mu)$ defined by

$$\nu R\nu' \Leftrightarrow \int_{\Delta} q d\nu(q) = \int_{\Delta} q d\nu'(q).$$

In fact, if $\Gamma(p)\neq\emptyset$, there is $\nu_p \in \Delta^{\sigma}(\mathcal{B}(\Delta),\mu)$ such that $\int_{\Delta} q d\nu_p(q) = p$, thus

$$[\nu_p] = \left\{\nu \in \Delta^{\sigma}(\mathcal{B}(\Delta),\mu) : \int_{\Delta} q d\nu(q) = p\right\} = \Gamma(p).$$

Conversely, for each $\nu' \in \Delta^{\sigma}(\mathcal{B}(\Delta),\mu)$, it is easy to show that $p' = \int_{\Delta} q d\nu'(q) \in \Delta$, thus

$$[\nu'] = \left\{\nu \in \Delta^{\sigma}(\mathcal{B}(\Delta),\mu) : \int_{\Delta} q d\nu(q) = p'\right\} = \Gamma(p')$$

and $\Gamma(p') \ni \nu'$ is not empty.

Finally, for each $f \in \mathcal{F}$,

$$\begin{aligned}
\phi^{-1}\left(\int_{\Delta} \phi\left(\int_S u(f)\,dp\right) d\mu(p)\right) &= \min_{p\in\Delta:\Gamma(p)\neq\emptyset}\left(\int_S u(f)\,dp + \min_{\nu'\in\Gamma(p)} I_{\int_S u(f)dp}(\nu' \parallel \mu)\right) \\
&= \min_{\nu\in\Delta^{\sigma}(\mathcal{B}(\Delta),\mu)}\left\{\int_{\Delta}\left(\int_S u(f)\,dq\right) d\nu(q) + \min_{\nu'\in[\nu]} I_{\int_{\Delta}\left(\int_S u(f)dq\right)d\nu(q)}(\nu' \parallel \mu)\right\} \\
&= \min_{\nu\in\Delta^{\sigma}(\mathcal{B}(\Delta),\mu)}\left\{\int_{\Delta}\left(\int_S u(f)\,dq\right) d\nu(q) + I_{\int_{\Delta}\left(\int_S u(f)dq\right)d\nu(q)}(\nu \parallel \mu)\right\}
\end{aligned}$$

as wanted. ■

Proof of Theorem 21. $C^* = \mathrm{cl}(\mathrm{dom}_{\Delta} G)$ follows from Theorem 10. Next, we show that $C^* = \mathrm{cl}(\mathrm{co}(\mathrm{supp}\mu))$ under the following weaker assumption on ϕ: for each $\alpha \in (0,1]$ and for each $a, b > 0$ there exists $n \in \mathbb{N}$ such that $\phi(0) > \alpha\phi(-an) + (1-\alpha)\phi(bn)$.

Since $\mu(\mathrm{supp}\mu) = 1$, then $\succsim$ is represented by

$$V(f) = \phi^{-1}\left(\int_{\mathrm{supp}\mu} \phi\left(\int u(f)\,dp\right) d\mu\right) \quad \forall f \in \mathcal{F}. \tag{83}$$

Consider the binary relation $\succsim^{**}$ defined by

$$f \succsim^{**} g \text{ if and only if } \int u(f)\,dp \geq \int u(g)\,dp \quad \forall p \in \mathrm{supp}\mu. \tag{84}$$

Then, using (84) and (83), it is immediate to check that $f \succsim^{**} g$ implies $f \succsim^{*} g$. By [25, Proposition A.1.], it follows $C^* \subseteq \mathrm{cl}(\mathrm{co}(\mathrm{supp}\mu))$.

Viceversa, suppose that $\mathrm{cl}(\mathrm{co}(\mathrm{supp}\mu)) \backslash C^* \neq \emptyset$. Then, since C^* is closed and convex, it cannot be the case that $\mathrm{supp}\mu \subseteq C^*$. Now, let $\bar{p} \in \mathrm{supp}\mu \backslash C^*$. Then, there exists $\psi \in B_0(\Sigma)$ and $a > 0$ such that

$$\int \psi d\bar{p} < -a < 0 < a \leq \int \psi dp \quad \forall p \in C^*. \tag{85}$$

Let $O = \left\{p \in \Delta : \int \psi dp < -a\right\}$. O is weak* open and contains $\bar{p}$, thus $O \cap \operatorname{supp}\mu \neq \emptyset$ and $\alpha = \mu(O) > 0$. Since $u(X) = \mathbb{R}$, define $\{f_n\}_{n\in\mathbb{N}} \subseteq \mathcal{F}$ to be such that $u(f_n) = n\psi$. Set $b = \max_{p\in\Delta} \int u(f_1)\, dp$. Notice that (85) implies that $b > 0$ and $\sup_{p\in O} \int u(f_1)\, dp \leq -a$. Then, for all $n \in \mathbb{N}$,

$$\begin{aligned}\int \phi\left(\int u(f_n)\, dp\right) d\mu &= \int_O \phi\left(\int u(f_n)\, dp\right) d\mu + \int_{O^c} \phi\left(\int u(f_n)\, dp\right) d\mu \\ &\leq \int_O \phi(-na)\, d\mu + \int_{O^c} \phi(nb)\, d\mu = \phi(-na)\,\alpha + (1-\alpha)\,\phi(nb)\,.\end{aligned}$$

The assumptions on ϕ guarantee that there exists $\bar{n} \in \mathbb{N}$ such that $\phi(-\bar{n}a)\,\alpha + (1-\alpha)\,\phi(\bar{n}b) < \phi(0)$. Take $x \in X$ such that $u(x) = 0$, then $\phi^{-1}\left(\int \phi\left(\int u(f_{\bar{n}})\, dp\right) d\mu\right) < 0$ implies $f_{\bar{n}} \prec x$. But, (85) implies $f_{\bar{n}} \succsim^* x$, hence $f_{\bar{n}} \succsim x$, a contradiction. ■

Proof of Corollary 22. Set $\phi(t) = -e^{-\theta t}$, then $\phi^{-1}(t) = -\theta^{-1}\log(-t)$, and

$$\phi^{-1}\left(\int_\Delta \phi(\langle p, \varphi\rangle)\, d\mu(p)\right) = -\frac{1}{\theta}\log\int_\Delta e^{-\theta\langle p,\varphi\rangle} d\mu(p)$$

for all $\varphi \in B_0(\Sigma)$; call this functional $I(\varphi)$. Let $(t, p) \in \mathbb{R} \times \Delta$. By (64),

$$\sup_{\varphi\in B_0(\Sigma):\langle p,\varphi\rangle\leq t} I(\varphi) = \phi^{-1}\left(\inf\left\{\inf_{k\geq 0}\left[tk - \int_\Delta \phi^*\left(k\frac{d\nu}{d\mu}\right) d\mu\right] : \nu \in \Gamma(p)\right\}\right)$$

while by Proposition 59, $\sup_{\varphi\in B_0(\Sigma):\langle p,\varphi\rangle\leq t} I(\varphi) = t + \frac{1}{\theta}\inf\{R(\nu \parallel \mu) : \nu \in \Gamma(p)\}$. Theorem 19 delivers the equivalence between (i) and (ii), while Proposition 12 that between (ii) and (iii).[35] ■

Proof of Theorem 23. By Proposition 58, the functional $I(\varphi) = \phi^{-1}\left(\int_\Delta \phi(\langle p, \varphi\rangle)\, d\mu(p)\right)$ for all $\varphi \in B_0(\Sigma)$, is translation invariant for all $\mu \in \Delta^\sigma(\mathcal{B}(\Delta))$ if and only if ϕ is CARA.

Next we show that for each given $\mu \in \Delta^\sigma(\mathcal{B}(\Delta))$, (u, ϕ, μ) represents a variational preference if and only if I is translation invariant.

Assume (u, ϕ, μ) represents a variational preference with variational representation (v, c). As observed in the proof of Theorem 19, I is well defined, quasiconcave, continuous, monotone, normalized, and $f \succsim g$ if and only if $I(u(f)) \geq I(u(g))$. But, by definition of variational representation, the functional $\bar{I}(\varphi) = \min_{p\in\Delta}(\langle p, \varphi\rangle + c(p))$ for all $\varphi \in B_0(\Sigma)$, (which is concave, continuous, monotone, normalized, and translation invariant) is such that $f \succsim g$ if and only if $\bar{I}(v(f)) \geq \bar{I}(v(g))$. But then, there are $\alpha > 0$ and $\beta \in \mathbb{R}$ such that $u = \alpha v + \beta$, and $(u, \alpha c)$ is a variational representation of $\succsim$. Then the functional $\tilde{I}(\varphi) = \min_{p\in\Delta}(\langle p, \varphi\rangle + \alpha c(p))$ for all $\varphi \in B_0(\Sigma)$, (which is concave, continuous, monotone,

[35] Notice that $\mu \in \Gamma\left(\int_\Delta q d\mu(q)\right)$, hence $\inf_{p\in\Delta}\left(\theta^{-1}\inf\{R(\gamma \parallel \mu) : \gamma \in \Gamma(p)\}\right) = 0$.

normalized, and translation invariant) is such that $f \succsim g$ if and only if $\tilde{I}(u(f)) \geq \tilde{I}(u(g))$. By Lemma 61, $I = \tilde{I}$ and I is translation invariant.

Conversely, if I is translation invariant, consider the preference $\succsim$ represented by (u, ϕ, μ). I is well defined, quasiconcave, continuous, monotone, normalized, and

$$f \succsim g \Longleftrightarrow \int_\Delta \phi\left(\int_S u(f)\,dp\right) d\mu(p) \geq \int_\Delta \phi\left(\int_S u(g)\,dp\right) d\mu(p) \Longleftrightarrow I(u(f)) \geq I(u(g)).$$

It is easy to check that $\succsim$ satisfies Axiom A.9 (on top of Axioms A.1-A.6), thus it is a variational preference. ■

Proof of Theorem 26. By Theorem 3, G is lower semicontinuous. Set, for all $\varphi \in B_0(\Sigma)$, $I(\varphi) = \min_{p\in\Delta} G\left(\int \varphi dp, p\right)$. Since (u, G) is a representation, $I : B_0(\Sigma) \to \mathbb{R}$ is normalized and $f \succsim g \Longleftrightarrow I(u(f)) \geq I(u(g))$. By Lemmas 61 and 62, $I : B_0(\Sigma) \to \mathbb{R}$ is normalized, monotone, quasiconcave, and uniformly continuous. Moreover, for all $(t, p) \in \mathbb{R} \times \Delta$,

$$G(t,p) = \sup_{f\in\mathcal{F}} \left\{ u(x_f) : \int u(f)\,dp \leq t \right\} = \sup_{f\in\mathcal{F}:\langle p, u(f)\rangle \leq t} I(u(f)) = \sup_{\varphi\in B_0(\Sigma):\langle p,\varphi\rangle\leq t} I(\varphi). \quad (86)$$

(i) implies (ii). For all $f \in \mathcal{F}$ and $\alpha \in (0, 1)$, A.10 implies that

$$f \sim x_f \Longrightarrow \alpha f + (1-\alpha)x_* \sim \alpha x_f + (1-\alpha)x_* \quad (87)$$

thus,[36] for all $\phi = u(f) \in B_0(\Sigma)$, $I(\alpha\phi) = I(\alpha u(f) + (1-\alpha)u(x_*)) = I(u(\alpha f + (1-\alpha)x_*))$ $= I(u(\alpha x_f + (1-\alpha)x_*)) = u(\alpha x_f + (1-\alpha)x_*) = \alpha u(x_f) + (1-\alpha)u(x_*) = \alpha u(x_f) = \alpha I(u(f)) = \alpha I(\phi)$. If $\alpha > 1$ we then have $I(\phi) = I(\alpha^{-1}\alpha\phi) = \alpha^{-1}I(\alpha\phi)$, and we conclude that $I : B_0(\Sigma) \to \mathbb{R}$ is positively homogeneous. By Theorem 45, this implies (ii).

(ii) implies (iii). Let $(t, p) \in \mathbb{R} \times \Delta$, by (ii)

$$G(t,p) = \begin{cases} \frac{t}{c_1(p)} & \text{if } t \geq 0 \text{ and } p \in C \\ \frac{t}{c_2(p)} & \text{if } t < 0 \text{ and } p \in C \\ \infty & \text{if } p \in \Delta \setminus C \end{cases}$$

$$= \begin{cases} \frac{t}{c_1(p)} = |t|\frac{1}{c_1(p)} & \text{if } t \geq 0 \text{ and } p \in C \\ \infty = t \times \infty = |t| \times \infty & \text{if } t \geq 0 \text{ and } p \in \Delta \setminus C \\ \frac{t}{c_2(p)} = (-t) \times \left(-\frac{1}{c_2(p)}\right) = |t| \times \left(-\frac{1}{c_2(p)}\right) & \text{if } t < 0 \text{ and } p \in C \\ \infty = -t \times \infty = |t| \times \infty & \text{if } t < 0 \text{ and } p \in \Delta \setminus C. \end{cases}$$

It suffices to set $\gamma(t) = |t|$ for all $t \in \mathbb{R}$,

$$d_1(p) = \begin{cases} \frac{1}{c_1(p)} & \text{if } p \in C \\ \infty & \text{if } p \in \Delta \setminus C \end{cases} \quad \text{and } d_2(p) = \begin{cases} -\frac{1}{c_2(p)} & \text{if } p \in C \\ \infty & \text{if } p \in \Delta \setminus C \end{cases}$$

[36] Notice that condition (87) is weaker than A.10, and it is sufficient to drive the result.

to obtain (iii).

(iii) implies (i). If $t > 0$, then $\gamma(t) > 0$ and, since $G \in \mathcal{G}(\mathbb{R} \times \Delta)$, $t = \inf_{p\in\Delta} G(t,p) = \inf_{p\in\Delta} \gamma(t) d_1(p) = \gamma(t) \inf_{p\in\Delta} d_1(p)$. Thus $\inf_{p\in\Delta} d_1(p) = a \in (0,\infty)$, and $\gamma(t) = a^{-1}t$. Analogously, if $t < 0$, then $\gamma(t) > 0$, and $t = \inf_{p\in\Delta} G(t,p) = \inf_{p\in\Delta} \gamma(t) d_2(p) = \gamma(t) \inf_{p\in\Delta} d_2(p)$, hence $\inf_{p\in\Delta} d_2(p) = -b$, with $b \in (0,\infty)$, and $\gamma(t) = -b^{-1}t$. Thus, for all $(t,p) \in \mathbb{R} \times \Delta$,

$$G(t,p) = \begin{cases} \gamma(t) d_1(p) & \text{if } t \geq 0 \text{ and } p \in \Delta \\ \gamma(t) d_2(p) & \text{if } t < 0 \text{ and } p \in \Delta \end{cases} = \begin{cases} t\frac{d_1(p)}{a} & \text{if } t \geq 0 \text{ and } p \in \Delta \\ t\frac{d_2(p)}{-b} & \text{if } t < 0 \text{ and } p \in \Delta \end{cases}$$

and $G(\alpha t, p) = \alpha G(t,p)$ for all $(t,p) \in \mathbb{R} \times \Delta$ and $\alpha > 0$. In turn, this implies I is positively homogeneous. Together with $u(x_*) = 0$, this allows to show that $\succsim$ satisfies Axiom A.10.■

Proof of Corollary 27. (i) implies (ii). Immediately descends from Theorem 26.

(ii) implies (i). Let, for all $(t,p) \in \mathbb{R} \times \Delta$,

$$G(t,p) = \begin{cases} \frac{t}{c_1(p)} & \text{if } t \geq 0 \text{ and } p \in C \\ \frac{t}{c_2(p)} & \text{if } t < 0 \text{ and } p \in C \\ \infty & \text{if } p \in \Delta \setminus C \end{cases}$$

and $I(\varphi) = \inf_{p\in\Delta} G(\langle p,\varphi\rangle, p)$ for all $\varphi \in B_0(\Sigma)$. By Lemma 44, I is finite, monotone, upper semicontinuous, positively homogeneous, quasiconcave

$$I(\varphi) = \min_{p\in C}\left(\frac{\langle p,\varphi\rangle^+}{c_1(p)} - \frac{\langle p,\varphi\rangle^-}{c_2(p)}\right) = \min_{p\in\Delta} G(\langle p,\varphi\rangle, p) \qquad \forall \varphi \in B_0(\Sigma) \tag{88}$$

and

$$\sup_{\psi\in B_0(\Sigma):\langle p,\psi\rangle\leq t} I(\psi) = G(t,p) \qquad \forall (t,p) \in \mathbb{R} \times \Delta. \tag{89}$$

Moreover, by (89) and Theorem 45, I is monotone, quasiconcave, uniformly continuous, positively homogeneous, and normalized. While, by (36) and (88), for all f and g in $\mathcal{F}$, $f \succsim g$ if and only if $I(u(f)) \geq I(u(g))$. By Lemmas 61, 62, and 63, $\succsim$ satisfies Axioms A.4-A.7, and it is easy to show that positive homogeneity guarantees that also A.10 holds.

In this case, by (89) and Theorem 41, (u,G) is a representation of $\succsim$ in the sense of Theorem 7. If Σ is a σ-algebra, then $\succsim$ satisfies Axiom A.8 if and only if there is $q \in \Delta^\sigma$ such that $G(\cdot,p) \equiv \infty$ for all $p \notin \Delta^\sigma(q)$, that is if and only if there is $q \in \Delta^\sigma$ such that $C \subseteq \Delta^\sigma(q)$.

Finally, if $v : X \to \mathbb{R}$ is affine and onto, with $v(x_*) = 0$, D is a nonempty, closed, and convex subset of Δ, and $d_1, d_2 : D \to [0,\infty]$ are functions such that the first concave and upper semicontinuous, with $0 < \inf_{p\in D} d_1(p) \leq \max_{p\in D} d_1(p) = 1$, the second convex and lower semicontinuous, with $\min_{p\in D} d_2(p) = 1$, and for all f and g in $\mathcal{F}$, $f \succsim g$ if and only if $\min_{p\in D}\left(\frac{\left(\int v(f)dp\right)^+}{d_1(p)} - \frac{\left(\int v(f)dp\right)^-}{d_2(p)}\right) \geq \min_{p\in D}\left(\frac{\left(\int v(g)dp\right)^+}{d_1(p)} - \frac{\left(\int v(g)dp\right)^-}{d_2(p)}\right)$. Then u and v represent $\succsim$ on X, therefore there is $\alpha > 0$ such that $v = \alpha u$. Thus

$$f \succsim g \Longleftrightarrow \min_{p\in D}\left(\frac{\left(\int u(f)\,dp\right)^+}{d_1(p)} - \frac{\left(\int u(f)\,dp\right)^-}{d_2(p)}\right) \geq \min_{p\in D}\left(\frac{\left(\int u(g)\,dp\right)^+}{d_1(p)} - \frac{\left(\int u(g)\,dp\right)^-}{d_2(p)}\right). \tag{90}$$

Set $J(\varphi) = \min_{p\in D}\left(\frac{\langle p,\varphi\rangle^+}{d_1(p)} - \frac{\langle p,\varphi\rangle^-}{d_2(p)}\right)$ for all $\varphi \in B_0(\Sigma)$. It can be shown, as we did in the proof that (ii) implies (i), that J is monotone, quasiconcave, uniformly continuous, positively homogeneous, and normalized, moreover, by (90), for all f and g in $\mathcal{F}$, $f \succsim g$ if and only if $J(u(f)) \geq J(u(g))$. By Lemma 61, $I = J$ and, by Lemma 44, $(C, c_1, c_2) = (D, d_1, d_2)$. ■

References

[1] J. Aczel, *Lectures on functional equations and their applications*, Academic Press, New York, 1966.

[2] C. D. Aliprantis and K. C. Border, *Infinite dimensional analysis*, 3rd ed., Springer Verlag, Berlin, 2006.

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

[4] K. Arrow, *Essays in the theory of risk-bearing*, North-Holland, Amsterdam, 1970.

[5] F. Bellini and M. Frittelli, On the existence of minimax martingale measures, *Mathematical Finance*, 12, 1–21, 2002.

[6] T. Bewley, Knightian decision theory: part I, *Decisions in Economics and Finance*, 25, 79–110, 2002.

[7] S. Cerreia-Vioglio, F. Maccheroni, M. Marinacci, and L. Montrucchio, Complete quasiconcave monotone duality, Carlo Alberto WP 80, 2008.

[8] S. Cerreia-Vioglio, F. Maccheroni, M. Marinacci, and L. Montrucchio, Probabilistic sophistication and uncertainty aversion, mimeo, 2009.

[9] S. Cerreia-Vioglio, F. Maccheroni, M. Marinacci, and L. Montrucchio, Risk measures: Rationality and diversification, Carlo Alberto WP 100, 2008.

[10] A. Chateauneuf and J. H. Faro, Ambiguity through confidence functions, mimeo, 2006.

[11] A. Chateauneuf, F. Maccheroni, M. Marinacci and J.-M. Tallon, Monotone continuous multiple priors, *Economic Theory*, 26, 973–982, 2005.

[12] K. M. Chong and N. M. Rice, Equimeasurable rearrangements of functions, *Queens Papers in Pure and Applied Mathematics*, 28, 1971.

[13] I. Csiszar, I-divergence geometry of probability distributions and minimization problems, *Annals of Probability*, 3, 146–158, 1975.

[14] G. Dal Maso, *An introduction to* Γ*-convergence*, Birkhäuser, Boston, 1993.

[15] G. Debreu, *Theory of value*, Yale University Press, 1959.

[16] B. de Finetti, Sulle stratificazioni convesse, *Annali di Matematica Pura e Applicata*, 30, 173–183, 1949.

[17] B. de Finetti, Sul concetto di media, *Giornale dell'Istituto Italiano degli Attuari*, 2, 369–396, 1931.

[18] W. E. Diewert, Duality approaches to microeconomic theory, in *Handbook of Mathematical Economics*, (K. J. Arrow and M. D. Intriligator, eds.), 4th edition, v. 2, pp. 535–599, Elsevier, Amsterdam, 1993.

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

[20] L. G. Epstein and M. Schneider, Recursive multiple-priors, *Journal of Economic Theory*, 113, 1–31, 2003.

[21] H. Ergin and F. Gul, A subjective theory of compound lotteries, mimeo, 2004.

[22] W. Fenchel, A remark on convex sets and polarity, *Meddelanden Lunds Universitets Matematiska Seminarium*, 82–89, 1952.

[23] M. Frittelli, Introduction to a theory of value coherent with the no-arbitrage principle, Finance and Stochastics, 4, 275–297, 2000.

[24] P. Gänssler, Compactness and sequential compactness in spaces of measures, *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 17, 124–146, 1971.

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

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

[27] P. Ghirardato and M. Siniscalchi, MoUSe: Monotonic, uncertainty-sensitive preferences, personal communication, 2009.

[28] I. Gilboa, F. Maccheroni, M. Marinacci, and D. Schmeidler, Objective and subjective rationality in a multiple prior model, mimeo, 2008.

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

[30] E. Hanany and P. Klibanoff, Updating ambiguity averse preferences, mimeo, 2007.

[31] L. P. Hansen, Beliefs, doubts, and learning: valuing macroeconomic risk, *American Economic Review*, 97, 1–30, 2007.

[32] L. P. Hansen and T. Sargent, Robust control and model uncertainty, *American Economic Review*, 91, 60–66, 2001.

[33] I. Karatzas, J. P. Lehoczky, S. E. Shreve, and G. L. Xu, Martingale and duality methods for utility maximization in an incomplete market, *SIAM Journal on Control and Optimization*, 29, 702–730, 1991.

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

[35] M. A. Krasnosel'sii and Y. B. Ruticjii, *Convex functions and Orlicz spaces*, Gordon and Breach Science Publishers, New York, 1961.

[36] F. Liese and I. Vajda, *Convex statistical distances*, Teubner, Leipzig, 1987.

[37] W. A. J. Luxemburg and A. C. Zaanen, *Riesz spaces*, vol. I, North-Holland, Amsterdam, 1971.

[38] F. Maccheroni, M. Marinacci, and A. Rustichini, Ambiguity aversion, robustness, and the variational representation of preferences, *Econometrica*, 74, 1447–1498, 2006.

[39] F. Maccheroni, M. Marinacci, and A. Rustichini, Dynamic variational preferences, *Journal of Economic Theory*, 128, 4–44, 2006.

[40] M. J. Machina and D. Schmeidler, A more robust definition of subjective probability, *Econometrica*, 60, 745–780, 1992.

[41] M. Marinacci and L. Montrucchio, Introduction to the mathematics of ambiguity, in *Uncertainty in Economic Theory*, (I. Gilboa, ed.), Routledge, New York, 2004.

[42] M. Marinacci and L. Montrucchio, On concavity and supermodularity, *Journal of Mathematical Analysis and Applications*, 344, 642–654, 2008.

[43] R. Nau, Uncertainty aversion with second-order utilities and probabilities, *Management Science*, 52, 136–145, 2004.

[44] J-P. Penot and M. Volle, On quasi-convex duality, *Mathematics of Operations Research,* 15, 597–625, 1990.

[45] L. Rigotti, C. Shannon, and T. Strzalecki, Subjective beliefs and ex-ante trade, *Econometrica*, 76, 1167–1190, 2008.

[46] R. T. Rockafellar, *Convex analysis*, Princeton University Press, Princeton, 1970.

[47] R. T. Rockafellar, Integrals which are convex functionals, *Pacific Journal of Mathematics*, 24, 525–539, 1968.

[48] R. T. Rockafellar, Integrals which are convex functionals, II, *Pacific Journal of Mathematics*, 39, 439–469, 1971.

[49] L. J. Savage, *The foundations of statistics*, Wiley, New York, 1954.

[50] W. Schachermayer, Optimal Investment in incomplete markets when wealth may become negative, *Annals of Applied Probability*, 11, 694–734, 2001.

[51] D. Schmeidler, Subjective probability and expected utility without additivity, *Econometrica*, 57, 571–587, 1989.

[52] K. Seo, Ambiguity and second-order belief, *Econometrica*, forthcoming.

[53] M. Siniscalchi, Dynamic choice under ambiguity, mimeo, 2006.

[54] T. Strzalecki, Axiomatic foundations of multiplier preferences, mimeo, 2007.

[55] K. Yosida, *Functional analysis*, Springer, New York, 1980.

[56] A. C. Zaanen, *Riesz spaces II*, North-Holland, Amsterdam, 1983.