THE CARLO ALBERTO NOTEBOOKS

Collegio Carlo Alberto

Uncertainty Averse Preferences

Simone Cerreia-Vioglio
Fabio Maccheroni
Massimo Marinacci
Luigi Montrucchio

Working Paper No. 77
September 2008
www.carloalberto.org

Uncertainty Averse Preferences[1]

Simone Cerreia-Vioglio[a] Fabio Maccheroni[b] Massimo Marinacci[c] Luigi 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 September 2008

[1]We thank Larry Epstein, Bart Lipman, Jawwad Noor, 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.

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 same time general and rich in structure. Many objective functions commonly used in applications are special cases of this representation.

JEL classification: D81

Keywords: Ambiguity Aversion; Games against Nature, Model Uncertainty; Smooth Ambiguity Preferences; Variational Preferences

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 in 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, 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 the 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.[4] 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 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 $\succsim$.

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.

[1] See, e.g., Rigotti, Shannon, and Strzalecki [40] and the references therein.

[2] See the classic discussions in Debreu [12, p. 101] and Schmeidler [45].

[3] Along, for example, the lines of Bewley [6]. See also the discussion in Gilboa, Maccheroni, Marinacci, and Schmeidler [23].

[4] We use the general term uncertainty – rather than a more specific term like, for example, ambiguity – because of the great generality of this class of preferences.

The function G plays a key role in the representation (1) and its properties are what gives (1) its rich structure. In particular, a noteworthy feature of (1) is the presence of expected utilities in the first argument, even though no independence assumption whatsoever is made on uncertain acts. Remarkably, expected utility thus already emerges in the representation of uncertainty averse preferences, and this confirms its prominent role in decision theory.

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.

1.1 Generality and Structure

The combination of generality and rich structure is the main feature of the representation (1). Thanks 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.[5]

All this can be seen in Section 4, 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 [34] and the smooth ambiguity preferences studied by Klibanoff, Marinacci, and Mukerji [30].[6]

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 10 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 [34] shows, the variational representation (2) includes as special cases the multiple priors model of Gilboa and Schmeidler [24] and the multiplier preferences of Hansen and Sargent ([28], [27]), which

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

[6] See Ergin and Gul [17], Nau [38], and Seo [46] for works related to [30].

can therefore be viewed as particular specifications of an additively separable uncertainty aversion index G.[7]

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. Theorem 16 shows that smooth preferences with concave ϕ correspond to the uncertainty aversion index given by

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

Here, $I_t(\cdot \parallel \mu)$ is a suitable statistical distance function, defined in (19), 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.

In the important exponential case $\phi(t) = -e^{-\theta t}$, Corollary 17 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 smooth preference functional (3) can thus be represented as

$$V(f) = \min_{p \in \Delta} \left\{ \int_{S} u(f(s))\, dp(s) + \frac{1}{\theta} \min_{\nu \in \Gamma(p)} R(\nu \parallel \mu) \right\}. \tag{5}$$

The preference functional (5) 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 18, shows that the overlap between these two classes of preferences is basically 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 preferences to consider are those characterized by multiplicatively separable uncertainty aversion indices. To further illustrate the flexibility of the representation (1), we carry out this exercise, which is related to the analysis of Chateauneuf and Faro [8], in Section 4.

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. In this case, f and p become, respectively, the strategies of the decision maker and of Nature.

[7] See Strzalecki (2007) for conditions on variational preferences that characterize multiplier preferences (see also Subsection 4.2.4 below).

[8] Recall that

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

As detailed in Section 3, the function $c : T \times \Delta \to (-\infty, \infty]$ such that

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

can be regarded as a a parametric cost function for Nature, where $c(t,p)$ is the cost for Nature to play p at value t of the parameter. Using this cost function, the objective function (1) can be written as

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

This is arguably the most general form of a game against Nature. Its special cases are determined by suitably specifying the parametric cost function c. For example, the variational representation (2) is characterized by a parametric cost function $c(t,p)$ that does not depend on t. The game theoretic interpretation of our setting thus generalizes the one discussed in [34] and [35] for variational preferences.

Notice how Nature's cost turns out to be parametrized by both players' strategies f and p through their expected utility $\int u(f)\, dp$. Also in the game interpretation, the appearance of expected utility at this level of generality (without any independence assumption) is remarkable. All this and more is discussed in Sections 3 (in particular in Subsection 3.4) and 4.

The analysis of this paper is static and its dynamic extension is a natural future research topic, along the lines of Epstein and Schneider [16] and Maccheroni, Marinacci, and Rustichini [35]. In this regard, it is also important to notice that Siniscalchi [47] and Hanany and Klibanoff [26] have recently studied in depth updating rules for uncertainty averse preferences; a natural direction of research is to see how their analysis can be read in terms of the representation (1).

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 the results on quasiconcave functions that are needed for our derivation. These quasiconcave methods are quite different from the concave duality methods that [34] 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. Section 4 studies some special classes of uncertainty averse preferences. Appendix C provides some more material on the statistical distance functions I_t, while Appendix D contains the proofs of the results.

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 $\succsim$. If $f \in \mathcal{F}$, 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 an 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 indifferently 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 $\mathrm{dom} G(\cdot, p) = \{t \in T : G(t, p) < \infty\}$. We denote by $\mathcal{E}(T \times \Delta)$ the set of functions in $\mathcal{H}(T \times \Delta)$ that have the following additional properties:

(vi) $\mathrm{dom} G(\cdot, p) \in \{\emptyset, T\}$ for all $p \in \Delta$,

(vii) $G(\cdot, p)$ are uniformly equicontinuous on T with respect to all $p \in \Delta$ such that $\mathrm{dom} G(\cdot, p) = T$.[10]

Property (vi) requires that the functions in $\mathcal{E}(T \times \Delta)$ be either real valued or constant at ∞; that is, either $\mathrm{dom} G(\cdot, p) = T$ or $\mathrm{dom} G(\cdot, p) = \emptyset$, respectively. Property (vii) requires that, 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

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

[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 $|t - t'| \leq \delta$ imply $|G(t, p) - G(t', p)| \leq \varepsilon$, for all $p \in \Delta$ such that $\mathrm{dom} G(\cdot, p) = T$.

from $B_0(\Sigma, T)$ to $[-\infty, \infty]$ is extended-valued continuous.[11] 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.*

A last piece of notation: we denote by $\mathcal{U}(X)$ the set of all nonconstant affine functions $u : X \to \mathbb{R}$.

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 (Uncertainty Aversion) *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 a convexity assumption that, as argued by Debreu [12] and Schmeidler [45], models a negative attitude toward the presence of uncertainty.

Definition 2 *A preference $\succsim$ is* uncertainty averse *if it satisfies axioms 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.

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 $\alpha \in (0, 1)$, $x \sim y$ implies $\alpha x + (1 - \alpha) z \sim \alpha y + (1 - \alpha) z$.*

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., [34, 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) *There are $x, y \in X$ such that, for each $\alpha \in (0, 1)$, there exist $z, z' \in X$ such that $\alpha z + (1 - \alpha) y \succ x \succ y \succ \alpha z' + (1 - \alpha) x$.*

[11] An ε-δ definition of linear continuity is given by Lemma 50 in Appendix A.

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*

$$f, g \in \mathcal{F} \text{ and } \frac{1}{2}f(s) + \frac{1}{2}y' \precsim \frac{1}{2}g(s) + \frac{1}{2}y \quad \forall s \in S \quad \Longrightarrow \quad \frac{1}{2}x_f + \frac{1}{2}z' \precsim \frac{1}{2}x_g + \frac{1}{2}z. \tag{6}$$

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 often a very convenient property because countably additive probabilities are much better behaved than probabilities that are merely finitely additive (see [9] and [34] 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 $E_1 \supseteq E_2 \supseteq \ldots$ and $\bigcap_{n\in\mathbb{N}} E_n = \emptyset$, then $f \succ g$ implies that there exists $n_0 \in \mathbb{N}$ such that $xE_{n_0}f \succ g$.*

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 $\mathcal{F}$. Then, the two following conditions are equivalent:*

(i) $\succsim$ is uncertainty averse 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 $\mathcal{F}$,

$$f \succsim g \Longleftrightarrow \inf_{p\in\Delta} G\left(\int u(f)\,dp, p\right) \geq \inf_{p\in\Delta} G\left(\int u(g)\,dp, p\right). \tag{7}$$

The function u is cardinally unique and, given u, there is a (unique) minimal $G^\star : u(X) \times \Delta \to (-\infty, \infty]$ in $\mathcal{G}(u(X) \times \Delta)$ satisfying (7), given by

$$G^\star(t, p) = \sup_{f\in\mathcal{F}} \left\{ u(x_f) : \int u(f)\,dp \leq t \right\} \qquad \forall (t, p) \in u(X) \times \Delta. \tag{8}$$

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 $u(X) \times \Delta$.[12]

Recall that x_f is a certainty equivalent of act f. Hence, thanks to (8) 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 $\int u(f)\,dp \leq t$.

[12] 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 $f \in \mathcal{F}$.

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

$$V(f) = \inf_{p \in \Delta} G^{\star}\left(\int u(f)\, dp, p\right), \qquad \forall f \in \mathcal{F}, \tag{9}$$

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

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 $\succsim$.[13] Such pairs have the following uniqueness property.

Proposition 4 *Let (u, G) be a 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 $\bar{G}(t, p) = \alpha G\left(\alpha^{-1}(t - \beta), p\right) + \beta$ 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 result shows that uniqueness holds when $\succsim$ satisfies A.6, that is, when $u(X) = \mathbb{R}$.

Proposition 5 *Let $\succsim$ be an uncertainty averse preference that satisfies A.4-A.6. Then, $G^{\star}$ defined in (8) is the unique lower semicontinuous $G \in \mathcal{G}(u(X) \times \Delta)$ for which (7) holds.*

3.3 Comparative Attitudes

Based on Ghirardato and Marinacci [22], 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 $x \in X$,

$$f \succsim_1 x \quad \Longrightarrow \quad f \succsim_2 x. \tag{10}$$

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 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.*

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

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

[13] In other words, 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 (7) and (8) hold.

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 the minimal uncertainty aversion of $\succsim$ 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 $t \in \mathbb{R}$. But 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 Games against Nature

As mentioned in the Introduction, our setting admits a game against Nature interpretation, where decision makers believe that they are 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 has to be suitably modified. For example, in Axiom A.3 the reason why decision makers prefer to randomize among indifferent acts is because this makes more costly for Nature (which has no control on the random device) to reply.

A key ingredient in this interpretation is the specification of a cost function of Nature. To this end, next we introduce parametric cost functions $c : T \times \Delta \to (-\infty,\infty]$, where $c(t,p)$ is the cost for Nature to play p at value t of the parameter.

Definition 7 *Given an interval T of the real line, a function $c : T \times \Delta \to (-\infty,\infty]$ is a* parametric cost function *if:*

(i) *c is nonnegative on $T \times \Delta$;*

(ii) *$c(t,\cdot)$ is quasiconvex on Δ for all $t \in T$;*

(iii) *$c(t,\cdot)$ is grounded, i.e., $\inf_{p\in\Delta} c(t,p) = 0$ for all $t \in T$.*

It is easy to check that if $G \in \mathcal{G}(T\times\Delta)$, then the difference $G(t,p) - t$ is indeed a parametric cost function. Therefore, in the equality

$$G(t,p) = t + c(t,p) \tag{11}$$

the function c is a parametric cost function. As a result, we can rewrite the representation (7) as

$$f \succsim g \iff \inf_{p\in\Delta}\left\{\int u(f)\, dp + c\left(\int u(f)\, dp, p\right)\right\} \geq \inf_{p\in\Delta}\left\{\int u(g)\, dp + c\left(\int u(g)\, dp, p\right)\right\}. \tag{12}$$

As observed in the Introduction, this is arguably the most general form of a game against Nature and special cases are determined by suitably specifying the parametric cost function c. In particular, this cost function is parametrized by both players' strategies f and p through their expected utility $\int u(f)\, dp$.

Representation (12) can be summarized by a pair (u,c), where c is a parametric cost function for Nature that corresponds, via (11), to an uncertainty averse representation. The comparative relation (10) can now be used to behaviorally pin down such cost functions. For, notice that, given a decision maker's act f, Nature can affect the relative likelihood of the act f outcomes by choosing a probabilistic model p, unless f is a constant act (in which case Nature has no power).

Hence, if $\succsim_2$ 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. The following version of Proposition 6 shows that in the representation (12) this translates into higher cost functions c for Nature.

Proposition 8 *Two preferences $\succsim_1$ and $\succsim_2$, with representations (u_1, c_1) and (u_2, c_2), satisfy (10) if and only if $u_1 \approx u_2$ and $c_1 \leq c_2$ (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 lower costs for her actions.

3.5 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 (even very weak form of independence actually ensure the Lipschitzianity of the representing preference functional).

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

Theorem 9 *Let $\succsim$ be a binary relation on $\mathcal{F}$. Then, the two 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 given by

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

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 actually suffice to have this equality sign, as well as the possibility of replacing Δ with $\Delta^\sigma(q)$ (provided Σ is a σ-algebra and $\succsim$ satisfies axiom A.8).

4 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.

4.1 Variational Preferences

We begin with the variational preferences of Maccheroni, Marinacci, and Rustichini [34]. 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 $\inf_{p \in \Delta} c(p) = 0$, such that

$$f \succsim g \Longleftrightarrow \min_{p \in \Delta} \left\{ \int u(f)\, dp + c(p) \right\} \geq \min_{p \in \Delta} \left\{ \int u(g)\, dp + c(p) \right\}, \tag{15}$$

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

As shown by [34], 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 [34].

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 (it is implied by A.9) and, setting

$$G(t, p) = t + c(p), \tag{16}$$

the pair (u, G) clearly represents $\succsim$ in the sense of (7). More is actually true:

Proposition 10 *Let $u : X \to \mathbb{R}$ be affine with $u(X) = \mathbb{R}$. If (u, c) is a variational representation of $\succsim$, then, setting*

$$G(t, p) = t + c(p) \qquad \forall (t, p) \in \mathbb{R} \times \Delta, \tag{17}$$

(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 aversion representations. Notice that in the game against Nature interpretation, the variational case corresponds to a parametric cost function $c(t, p)$ for Nature that does not depend on t.

4.2 Smooth Ambiguity Preferences

The smooth ambiguity preferences studied by Klibanoff, Marinacci, and Mukerji [30] provide another example of uncertainty averse preferences. In this subsection we will study 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

$$f \succsim g \Longleftrightarrow \int_{\Delta} \phi \left(\int_S u(f(s))\, dp(s) \right) d\mu(p) \geq \int_{\Delta} \phi \left(\int_S u(g(s))\, dp(s) \right) d\mu(p), \tag{18}$$

for all $f, g \in \mathcal{F}$.[14]

As in standard statistical decision theory, the first-order probabilities p are possible models that govern states' realizations, while the second-order probabilities μ are priors on such models. As discussed by [30], 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.

Throughout the paper we will consider the ϕ concave case. In order to establish the uncertainty averse representation of these smooth preferences, we need to introduce a family of statistical distance functions.[15]

4.2.1 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

$$I_t(\nu \parallel \mu) = \phi^{-1}\left(\inf_{k\geq 0}\left[kt - \int \phi^*\left(k\frac{d\nu}{d\mu}\right) d\mu\right]\right) - t. \tag{19}$$

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

Proposition 11 *For all $t \in \mathbb{R}$,*

(i) $I_t(\mu \parallel \mu) = 0$*;*

(ii) $I_t(\nu \parallel \mu) \geq 0$ *for each* $\nu \in \Delta^\sigma(\mathcal{B}(\Delta), \mu)$*;*

(iii) $I_t(\cdot \parallel \mu)$ *is quasiconvex;*

(iv) $I_t(\cdot \parallel \mu)$ *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* $c \in \mathbb{R}$.

Example 12 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 15 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 [19] 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 13 *Suppose Σ is not trivial. Then, given two strictly increasing and concave functions $\phi_1, \phi_2 : \mathbb{R} \to \mathbb{R}$, the following conditions are equivalent:*

[14] In richer settings (whose specification is beyond the scope of this paper), Ergin and Gul [17], Klibanoff, Marinacci, and Mukerji [30], Nau [38], and Seo [46] provide behavioral conditions that underlie the representation (18). 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 (65) in Appendix B.

[15] See [32] for a thorough study of statistical distance functions.

(i) ϕ_1 *is more concave than* ϕ_2;[16]

(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 $I^1 = I^2$. 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 $\phi'(0) = 1$.

Definition 14 *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 15 *If* ϕ *is order Orlicz, then*

$$I_t(\nu \parallel \mu) = \phi^{-1}\left(\int (\phi \circ \psi)\left(k(\nu)\frac{d\nu}{d\mu}\right)d\mu\right) - t, \qquad \forall t \in \mathbb{R}, \nu \in \mathrm{dom} I_t(\cdot \parallel \mu) \tag{20}$$

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

$$\int \psi\left(k\frac{d\nu}{d\mu}\right)d\nu = t. \tag{21}$$

In other words, when ϕ is order Orlicz, the index I_t can be computed in two stages. First, $k(\nu)$ is determined via (21), and then it is used to determine I_t via (20). This procedure is known (see [29]), our contribution is to identify a class of functions in which it works (see also [44]).

4.2.2 Uncertainty Averse Representation

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

$$p(A) = \int_\Delta q(A)\, d\nu(q) \qquad \forall A \in \mathcal{B}(\Delta).$$

Theorem 16 *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* $(t, p) \in \mathbb{R} \times \Delta$,

$$G(t, p) = t + \min_{\nu \in \Gamma(p)} I_t(\nu \parallel \mu), \tag{22}$$

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* $\Gamma(p) = \emptyset$.

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

The important part of Theorem 16 is (22), which provides an explicit formula for the uncertainty aversion index G in the smooth case.

To interpret this formula, first observe that 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.

When the support of μ is finite, say $\text{supp}(\mu) = \{q_1, ..., q_n\}$, there is at most one second-order probability ν with this property provided the first-order probabilities in $\text{supp}(\mu)$ are linearly independent. In fact, 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

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

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 $\text{supp}(\mu)$. The set $\Gamma(p)$ is nonempty if and only if p belongs to the convex hull of the support of μ; that is, $\Gamma(p) \neq \emptyset$ if and only if $p \in \text{co}(\text{supp}(\mu))$. When this happens, $\Gamma(p)$ is a singleton if the probabilities in $\text{supp}(\mu)$ are linearly independent. Thus, the nonsingleton nature of $\Gamma(p)$ reflects a linear dependence of the probabilities in $\text{supp}(\mu)$.

In view of this decision theoretic interpretation of $\Gamma(p)$, we can say that $G(t, p)$ is determined in formula (22) by evaluating all second-order probabilities ν in $\Gamma(p)$ through the distance $I_t(\nu \parallel \mu)$ with respect to μ. The least distant one is then selected. Probabilistically, the term $G(t, p) - t$, that is, $\min_{\nu\in\Gamma(p)} I_t(\nu \parallel \mu)$, is called the I_t distance of μ from $\Gamma(p)$ and an element of $\Gamma(p)$ where the minimum is achieved is called projection of μ on $\Gamma(p)$ (see Csiszar [10]).

Summing up, by Theorem 16,

$$\phi^{-1}\left(\int_\Delta \phi\left(\int_S u(f)\,dp\right) d\mu(p)\right) = \min_{p\in\Delta}\left(\int_S u(f)\,dp + \min_{\nu\in\Gamma(p)} I_{\int_S u(f)dp}(\nu \parallel \mu)\right) \tag{23}$$

for all $f \in \mathcal{F}$. Here the game against Nature interpretation is especially stark. In fact, the parametric cost function of Nature $c : \mathbb{R} \times \Delta \to (-\infty, \infty]$ is given by

$$c(t, p) = \min_{\nu\in\Gamma(p)} I_t(\nu \parallel \mu),$$

that is, by the I_t distance of μ from $\Gamma(p)$.

4.2.3 Exponential Case and Overlap

Consider the important exponential case $\phi(t) = -e^{-\theta t}$, which corresponds to constant ambiguity aversion (see [30, p. 1866]). In this case we have the following version of Theorem 16, where $I_t(\cdot\|\mu)$ reduces to the relative entropy $R(\cdot\|\mu)$.

Corollary 17 *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*

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

(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 (23) becomes:

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

Corollary 17 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 18 *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.*

4.2.4 Quasi-Arithmetic Representation and Multiplier Preferences

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

$$V(f) = \phi^{-1}\left(\int_S (\phi \circ u)(f) \, dq\right), \qquad \forall f \in \mathcal{F}, \tag{24}$$

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.[17] We call (u, ϕ, q) a *quasi-arithmetic representation* of $\succsim$ and we refer the interested reader to Strzalecki (2007) for a recent discussion of this setting.

When ϕ is the negative exponential $-e^{-\theta t}$, the representation (24) takes the variational form

$$V(f) = \min_{p \in \Delta^\sigma(q)} \left\{ \int_S u(f) \, dp + R(p \parallel q) \right\}, \tag{25}$$

with the relative entropy $R(p \parallel q)$ as cost function. This variational representation corresponds to the Hansen and Sargent multiplier preferences ([28], [27]). In particular, Strzalecki (2007) provided behavioral conditions on variational preferences that characterize (25).

When ϕ is a general concave function, not necessarily exponential, the quasi-arithmetic representation (24) 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 (24), thus generalizing its variational representation (25) obtained for the ϕ exponential case.[18]

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* $q \in \Delta^\sigma$ *a probability measure on* S*. The following conditions are equivalent:*

[17] It can be checked that a preference $\succsim$ satisfies Savage's axioms P1-P6 and Axioms A.4, A.5, and A.8 if and only if it can be represented by (24) with a nonatomic q. Moreover, $\succsim$ also satisfies Axioms A.3, A.6, and A.9 if and only if ϕ is CARA (these observations have been made jointly with Larry Epstein).

[18] We omit the proof of this result because it is essentially an elementary version of the more complicated Theorem 16 and Corollary 17. Similarly, we omit the proof of Proposition 20, which is a simpler version of that of Theorem 18.

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

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

$$G(t, p) = \begin{cases} t + I_t(p \parallel q) & \text{if } p \in \Delta^\sigma(q), \\ \infty & \text{else.} \end{cases} \tag{26}$$

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 (26), we thus have

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

for all $f \in \mathcal{F}$. In the game against Nature interpretation, this means that Nature's parametric cost function $c(t, p)$ is given by the statistical distance $I_t(p \parallel q)$.

Finally, a result parallel to Theorem 18 holds here.

Proposition 20 *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 (25) is basically the overlap between variational and quasi-arithmetic representations.

4.3 The Homothetic Case

Proposition 10 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 (2006), 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).

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 [24] (see the discussion in [34, 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 21 *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 $(t,p) \in \mathbb{R}\times\Delta$,

$$G(t,p) = \begin{cases} \frac{t}{c_1(p)} & \text{if } t \geq 0 \text{ and } p \in C \\ \frac{t}{c_2(p)} & \text{if } t < 0 \text{ and } p \in C \\ \infty & \text{if } p \in \Delta \setminus C \end{cases} \tag{27}$$

(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 $0 \cdot \infty = \infty$.

By Theorem 21, we have the following representation result.

Corollary 22 *Let* $\succsim$ *be a binary relation on* $\mathcal{F}$*. Then, the two 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 21 such that, for all* f *and* g *in* $\mathcal{F}$, $f \succsim g$ *if and only if*

$$\min_{p\in C}\left(\frac{\left(\int u(f)\,dp\right)^+}{c_1(p)} - \frac{\left(\int u(f)\,dp\right)^-}{c_2(p)}\right) \geq \min_{p\in C}\left(\frac{\left(\int u(g)\,dp\right)^+}{c_1(p)} - \frac{\left(\int u(g)\,dp\right)^-}{c_2(p)}\right). \tag{28}$$

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

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* $C \subseteq \Delta^\sigma(q)$.

For example, if $f(s), g(s) \succsim x_*$ for all $s \in S$, then (28) 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 (2006), who assume the existence of a worst outcome with respect to which A.10 holds.

We close with couple of remarks. First notice that c_2 can take value ∞. In particular, we can have $c_2(p) = \infty$ for all $p \in C$. In this case $G(t,p) = 0$ for all $t < 0$ and $p \in C$, and (28) becomes:

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

Second, we already observed that a preference satisfies the Certainty Independence axiom of Gilboa and Schmeidler [24] 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 (17) and (27), an index G is both variational and homothetic if:

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

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

$$f \succsim g \Longleftrightarrow \min_{p \in C} \int u(f)\, dp \geq \min_{p \in C} \int u(g)\, dp, \tag{29}$$

which is the multiple priors criterion. Notice that for fixed u and C, by Proposition 6, the agent using criterion (29) is the most uncertainty averse of those using criterion (28).

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.[19]

Notation 23 *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 [33] 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

Assumption 1 $(X, \|\cdot\|, \geq)$ *is a normed Riesz space with order unit e and $\|\cdot\|$ is its supnorm, i.e.*

$$\|x\| = \inf\{\alpha \in \mathbb{R} : |x| \leq \alpha e\} \qquad \forall x \in X.$$

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 have 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., [33, 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

$$[-e, e] = \{x \in X : -e \leq x \leq e\}. \tag{30}$$

Denoting by X_+ and X_- is the positive and negative cones in X, then the positive and negative unit balls are

$$[-e, e] \cap X_+ = [0, e] \text{ and } [-e, e] \cap X_- = [-e, 0].$$

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

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

[19] For the sake of completeness, in this version of the paper we report the proofs of almost all the formal statements with the following mark-up: an asterisk "*" for those results that are special cases of those in [7], a pound "#" for standard verifications.

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 (30), $\|\xi\| = \langle \xi, e\rangle$ 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.[20] Evenly convex sets are convex, and intersections of evenly convex sets are evenly convex.

Lemma 24 *A set C is evenly convex if and only if for all $\bar{x} \notin C$ there is $\bar{\xi} \in X^*$ such that $\left\langle \bar{\xi}, \bar{x}\right\rangle < \left\langle \bar{\xi}, x\right\rangle$ for all $x \in C$.*

By standard separation results, both open convex sets and closed convex sets are then evenly convex.

Lemma 25 *For every interval T of $\mathbb{R}$, $X(T)$ is evenly convex.*

Proof.# If T is a closed half line, then $X(T)$ is of the form $z \pm X_+$ for some $z \in \langle e\rangle$, and it is closed and convex.

If T is a open half line, then $X(T)$ is the interior of a set of the form $z \pm X_+$ for some $z \in \langle e\rangle$, and it is open and convex.

Else, there exist two half lines T' and T'' such that $T = T' \cap T''$, in this case

$$\begin{aligned} X(T) = X(T' \cap T'') &= \{x \in X : [\operatorname{ess\,inf}(x), \operatorname{ess\,sup}(x)] \subseteq T' \cap T''\} \\ &= X(T') \cap X(T''), \end{aligned}$$

and $X(T)$ is evenly convex being an intersection of evenly convex sets. ■

Functions

If $Y \subseteq X$, $g : Y \to [-\infty, \infty]$, and $a \in [-\infty, \infty]$, we set

$$\{g \geq a\} = \{y \in Y : g(y) \geq a\},$$

$\{g > a\}$, $\{g \leq a\}$, and $\{g < a\}$ are defined in the same way.

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

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

[20] 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 [18].

- *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.[21]

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:

- *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 (see [11, Ch. 3])

$$g^+(x) = \inf_{U \in \mathcal{N}_x} \sup_{y \in U} g(y), \qquad \forall x \in X;$$

and hence

$$\{g^+ \geq \alpha\} = \bigcap_{\beta < \alpha} \overline{\{g > \beta\}}, \qquad \forall \alpha \in \mathbb{R}. \tag{31}$$

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

Lemma 26 *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 $n \in \mathbb{N}$.*[22]

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 $V_{n_U} \subseteq U$,[23] and we also have

$$\sup_{y \in U} g(y) \geq \sup_{y \in V_{n_U}} g(y) \geq \inf_n \sup_{y \in V_n} g(y).$$

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

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

[23] 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 $[x - e_n, x + e_n] \subseteq [x - \delta e, x + \delta e] \subseteq U$.

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 $g\left(x + n^{-1}e\right) \leq g\left(z + n^{-1}e\right)$ for all $n \in \mathbb{N}$, whence

$$g^+(x) = \inf_n g\left(x + n^{-1}e\right) \leq \inf_n g\left(z + n^{-1}e\right) = g^+(z),$$

thus g^+ is monotone.

Finally, if g is quasiconcave, (31) 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 (see [11, Ch. 3])

$$g^-(x) = \sup_{U\in\mathcal{N}_x} \inf_{y\in U} g(y), \qquad \forall x \in X;$$

and hence

$$\left\{g^- > \alpha\right\} = \bigcup_{\beta>\alpha} \{g > \beta\}^\circ, \qquad \forall \alpha \in \mathbb{R}. \tag{32}$$

Moreover, g^- is the greatest lower semicontinuous function on X that is pointwise dominated by g.

Lemma 27 *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.*

Proof.# Let $x \in X$. For each $n \geq 1$, set $V_n = [x_n, 2x - x_n] = [x - e_n, x + e_n]$, where $e_n = x - x_n$ 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, $\sup_{U\in\mathcal{N}_x}\inf_{y\in U} g(y) \geq \sup_n \inf_{y\in V_n} g(y)$. Moreover, since $e_n \to 0$, for each $U \in \mathcal{N}_x$ there is $n_U \in \mathbb{N}$ such that $V_{n_U} \subseteq U$,[24] and we also have

$$\inf_{y\in U} g(y) \leq \inf_{y\in V_{n_U}} g(y) \leq \sup_n \inf_{y\in V_n} g(y).$$

Then $\sup_{U\in\mathcal{N}_x}\inf_{y\in U} g(y) \leq \sup_n \inf_{y\in V_n} g(y)$, and $g^-(x) = \sup_n \inf_{y\in V_n} g(y)$. By monotonicity of g, $\inf_{y\in V_n} g(y) = g(x_n)$ and $g^-(x) = \sup_n g(x_n)$.

If $z \in X$ and $x \leq z$, then $g\left(x - n^{-1}e\right) \leq g\left(z - n^{-1}e\right)$ for all $n \in \mathbb{N}$, whence

$$g^-(x) = \sup_n g\left(x - n^{-1}e\right) \leq \sup_n g\left(z - n^{-1}e\right) = g^-(z),$$

thus g^- is monotone.

Finally, if g is quasiconcave, then for all $\forall\alpha \in \mathbb{R}$ and $\beta > \alpha$ the sets $\{g > \beta\}$ and $\{g > \beta\}^\circ$ are convex. Moreover, if $x, y \in \bigcup_{\beta>\alpha} \{g > \beta\}^\circ$ there are $\beta_x, \beta_y > \alpha$ such that $x \in \{g > \beta_x\}^\circ$ and $y \in \{g > \beta_y\}^\circ$, wlog $\beta_x > \beta_y$, then $\{g > \beta_x\} \subseteq \{g > \beta_y\}$, $\{g > \beta_x\}^\circ \subseteq \{g > \beta_y\}^\circ$, and $x, y \in \{g > \beta_y\}^\circ$. Therefore, since $\{g > \beta_y\}^\circ$ is convex, all the convex combinations of x and y belong to $\{g > \beta_y\}^\circ \subseteq \bigcup_{\beta>\alpha} \{g > \beta\}^\circ$. Conclude that $\bigcup_{\beta>\alpha} \{g > \beta\}^\circ$ is convex and, by (32), g^- is quasiconcave. ■

[24] 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 $[x - e_n, x + e_n] \subseteq [x - \delta e, x + \delta e] \subseteq U$.

A.1.2 Two Key Auxiliary Functions

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

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

and

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

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

These two functions, which will play a key role in what follows, can take values on $[-\infty, \infty]$. For our analysis, the set where they can take on value ∞ is more relevant than that where they take on value $-\infty$. Hence, throughout the appendix we set $\operatorname{dom}(g_{\xi}) = \{g_{\xi} < \infty\}$ and $\operatorname{dom}(G_{\xi}) = \{G_{\xi} < \infty\}$.

The function G_{ξ} is monotone and dominates g_{ξ}. In fact, $G_{\xi}(t) = \sup_{k \leq t} g_{\xi}(k)$. Moreover:

(i) $g_{\alpha\xi}(\alpha t) = g_{\xi}(t)$ for all $\alpha \in \mathbb{R} \setminus \{0\}$;

(ii) $G_{\alpha\xi}(\alpha t) = G_{\xi}(t)$ for all $\alpha > 0$.

Denote by g_{ξ}^{+} and G_{ξ}^{+} the upper semicontinuous envelopes of g_{ξ} and G_{ξ}, respectively. In particular, by Lemma 26, $G_{\xi}^{+}(t) = \inf \{G_{\xi}(t') : t' > t\}$ since G_{ξ} is monotone.

The next lemmas give some basic properties of the function G_{ξ} (the proofs, when omitted, can be found in [7]).

Lemma 28 *For any function $g : Y \to [-\infty, \infty]$, the map $(t, \xi) \mapsto G_{\xi}(t)$ is quasiconvex over $\mathbb{R} \times \Delta$. Moreover,*

$$\lim_{t \to \infty} G_{\xi}(t) = \sup_{\zeta \in \Delta} \sup_{t \in \mathbb{R}} G_{\zeta}(t) = \sup_{x \in Y} g(x), \qquad \forall \xi \in \Delta.$$

Proof.* Let $(t_1, \xi_1), (t_2, \xi_2) \in \mathbb{R} \times \Delta$ and $\alpha \in (0, 1)$. Consider the point (t', ξ'), with $t' = \alpha t_1 + (1 - \alpha) t_2$ and $\xi' = \alpha \xi_1 + (1 - \alpha) \xi_2$. We have

$$\{x \in Y : \langle \xi', x \rangle \leq t'\} \subseteq \{x \in Y : \langle \xi_1, x \rangle \leq t_1\} \cup \{x \in Y : \langle \xi_2, x \rangle \leq t_2\}, \tag{33}$$

which implies $G_{\xi'}(t') \leq \max \{G_{\xi_1}(t_1), G_{\xi_2}(t_2)\}$, as desired.

Moreover, $G_{\zeta}(t) \leq \sup_{x \in Y} g(x)$ for all $t \in \mathbb{R}$ and all $\zeta \in \Delta$, so that $\sup_{\zeta \in \Delta} \sup_{t \in \mathbb{R}} G_{\zeta}(t) \leq \sup_{x \in Y} g(x)$. Similarly, $g(y) \leq G_{\xi}(\langle \xi, y \rangle)$ for all $y \in Y$ and all $\xi \in \Delta$.

There exists a sequence $\{x_n\} \in Y$ such that $g(x_n) \uparrow \sup_{x \in Y} g(x)$. Since $t \mapsto G_{\xi}(t)$ is monotone, we have $g(x_n) \leq G_{\xi}(\langle \xi, x_n \rangle) \leq \lim_{t \to \infty} G_{\xi}(t)$ for all $n \in \mathbb{N}$. Hence,

$$\sup_{x \in Y} g(x) = \lim_{n} g(x_n) \leq \lim_{t \to \infty} G_{\xi}(t) \leq \sup_{\zeta \in \Delta} \sup_{t \in \mathbb{R}} G_{\zeta}(t) \leq \sup_{x \in Y} g(x)$$

as desired. ■

Lemma 29 *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 30 *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 31 *If $g : X \to [-\infty, \infty]$ is monotone, then $G_\xi = g_\xi$ for all $\xi \in \Delta$.*

Proof.* Clearly, $g_\xi(t) \leq G_\xi(t)$ for all $(t, \xi) \in \mathbb{R} \times \Delta$. Suppose, by contradiction that $g_\xi(t) < G_\xi(t)$ for some $\xi \in \Delta$ and $t \in \mathbb{R}$. This implies that

$$\begin{aligned} \sup\{g(x) : x \in X \text{ and } \langle \xi, x \rangle = t\} &< \sup\{g(x) : x \in X \text{ and } \langle \xi, x \rangle \leq t\} \\ &= \sup\{g(x) : x \in X \text{ and } \langle \xi, x \rangle = t\} \vee \sup\{g(x) : x \in X \text{ and } \langle \xi, x \rangle < t\} \end{aligned}$$

and

$$\sup\{g(x) : x \in X \text{ and } \langle \xi, x \rangle = t\} < \sup\{g(x) : x \in X \text{ and } \langle \xi, x \rangle < t\}.$$

Therefore there exists a point $\bar{x} \in X$ for which $g_\xi(t) < g(\bar{x}) \leq G_\xi(t)$ and $\langle \xi, \bar{x} \rangle < t$. But $\langle \xi, \bar{x} + \alpha e \rangle = t$, for $\alpha = t - \langle \xi, \bar{x} \rangle > 0$. Hence, $g(\bar{x}) \leq g(\bar{x} + \alpha e) \leq g_\xi(t)$ that leads to a contradiction. ■

Lemma 32 *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, $G_\xi(t) = \varphi(H_\xi(t))$ and $g_\xi(t) = \varphi(h_\xi(t))$ for all $(\xi, t) \in \mathbb{R} \times \Delta$.*

Proof.# We prove that, in general

$$\sup_{x \in C} g(x) = \varphi\left(\sup_{x \in C} h(x)\right)$$

for each nonempty set C, $h : C \to [-\infty, \infty]$, $\varphi : \overline{h(C)} \to [-\infty, \infty]$ extended-valued continuous and monotone, and $g = \varphi \circ h$.

It is then sufficient to observe that, for all $(\xi, t) \in \mathbb{R} \times \Delta$, the sets $\{x \in X : \langle \xi, x \rangle \leq t\}$ and $\{x \in X : \langle \xi, x \rangle = t\}$ are not empty.

Since C is not empty, there exists sequence x_n in C such that $h(x_n) \to \sup_{x \in C} h(x)$. Therefore $H = \sup_{x \in C} h(x) \in \overline{h(C)}$. By definition $H \geq h(x)$ for all $x \in C$ and monotonicity of φ implies $\varphi(H) \geq \varphi(h(x)) = g(x)$ for all $x \in C$. That is, $\varphi(H)$ is an upper bound for g on C. Assume, per contra, there exists an upper bound m for g on C such that $m < \varphi(H)$, then $\varphi(h(x_n)) \to \varphi(H)$ and eventually $m < \varphi(h(x_n)) = g(x_n)$, a contradiction. ■

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 functions $g_\xi(t)$ and $G_\xi(t)$ 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 for the function g_ξ can be found in de Finetti [13, p. 178], while a closely related general formulation can be found in [39, Theorem 2.6]. Notice that versions of this result play an important role in microeconomic duality theory (see, e.g., Diewert [15]).

Theorem 33 *A function $g : X \to [-\infty, +\infty]$ is evenly quasiconcave and monotone if and only if*

$$g(x) = \inf_{\xi\in\Delta} G_\xi(\langle \xi, x\rangle) = \inf_{\xi\in\Delta} g_\xi(\langle \xi, x\rangle), \qquad \forall x \in X. \tag{34}$$

Moreover:

(i) If g is lower semicontinuous, then the infima in (34) are attained for all $x \in X$.

(ii) If g is upper semicontinuous, then G_ξ and g_ξ in (34) can be replaced with G_ξ^+ and g_ξ^+, respectively.

Proof.* First, by Lemma 31, $G_\xi(\langle \xi, x\rangle) = g_\xi(\langle \xi, x\rangle)$, for all $x \in X$ and $\xi \in \Delta$. Therefore it suffices to prove the statement only for the functions G_ξ.

"If" Suppose g is evenly quasiconcave and monotone.

If $g \equiv \pm\infty$, then

$$G_\xi(t) = \sup\{g(y) : y \in X \text{ and } \langle \xi, y\rangle \le t\} = \pm\infty$$

for all $(t, \xi) \in \mathbb{R}\times\Delta$, and the result is trivial.

Else, since

$$g(x) \le G_\xi(\langle \xi, x\rangle), \qquad \forall x \in X,\ \forall \xi \in \Delta, \tag{35}$$

then

$$g(x) \le \inf_{\xi\in\Delta} G_\xi(\langle \xi, x\rangle), \qquad \forall x \in X. \tag{36}$$

If $\bar{x}$ is a global maximum for g on X, equality holds in (35), and so in (36). Assume that $\bar{x} \in X$ is not a global maximum.

Let $r \in \mathbb{R}$ be such that $\{g \ge r\} \ne \emptyset$ and $\bar{x} \notin \{g \ge r\}$. Since the latter set is evenly convex, by Lemma 24, there exists $\bar{\xi} \in X^*$ such that $\langle \bar{\xi}, \bar{x}\rangle < \langle \bar{\xi}, x\rangle$ for all $x \in \{g \ge r\}$. Clearly, $\bar{\xi} \ne 0$ hence $\langle \bar{\xi}/\|\bar{\xi}\|, \bar{x}\rangle < \langle \bar{\xi}/\|\bar{\xi}\|, x\rangle$ for all $x \in \{g \ge r\}$. Wlog $\bar{\xi}$ belongs to the unit ball in X^*. Next we show that $\bar{\xi}$ is positive, thus (wlog) $\bar{\xi} \in \Delta$. Let $z \in X_+$ and take $y \in \{g \ge r\}$. Notice that, by monotonicity, $y + nz \in \{g \ge r\}$ for all $n \in \mathbb{N}$, and so $\langle \bar{\xi}, \bar{x}\rangle < \langle \bar{\xi}, y\rangle + n\langle \bar{\xi}, z\rangle$. Then, $\langle \bar{\xi}, z\rangle > n^{-1}\left(\langle \bar{\xi}, \bar{x}\rangle - \langle \bar{\xi}, y\rangle\right)$ for all $n \in \mathbb{N}$, which implies $\langle \bar{\xi}, z\rangle \ge 0$, as desired. We have shown the following:

Fact For all $r \in \mathbb{R}$ such that $\{g \ge r\} \ne \emptyset$ and $\bar{x} \notin \{g \ge r\}$, there exists $\bar{\xi} = \bar{\xi}_r \in \Delta$ such that $\langle \bar{\xi}, \bar{x}\rangle < \langle \bar{\xi}, x\rangle$ for all $x \in \{g \ge r\}$.

Case 1: Suppose $g(\bar{x}) \in \mathbb{R}$. Since $\bar{x}$ is not a global maximum, there is $\bar{\varepsilon} > 0$ such that $\{g \ge g(\bar{x}) + \varepsilon\} \ne \emptyset$ for all $\varepsilon \in (0, \bar{\varepsilon}]$. For all such ε, $\bar{x} \notin \{g \ge g(\bar{x}) + \varepsilon\} \ne \emptyset$. Then there exists $\bar{\xi} = \bar{\xi}_\varepsilon \in \Delta$ such that $\langle \bar{\xi}, \bar{x}\rangle < \langle \bar{\xi}, x\rangle$ for all $x \in \{g \ge g(\bar{x}) + \varepsilon\}$. That is, $\{g \ge g(\bar{x}) + \varepsilon\} \subseteq \{\bar{\xi} > \langle \bar{\xi}, \bar{x}\rangle\}$ and $\{\bar{\xi} \le \langle \bar{\xi}, \bar{x}\rangle\} \subseteq \{g < g(\bar{x}) + \varepsilon\}$. Thus, $G_{\bar{\xi}}(\langle \bar{\xi}, \bar{x}\rangle) \le g(\bar{x}) + \varepsilon$ and

$$g(\bar{x}) \le \inf_{\xi\in\Delta} G_\xi(\langle \xi, \bar{x}\rangle) \le G_{\bar{\xi}}(\langle \bar{\xi}, \bar{x}\rangle) \le g(\bar{x}) + \varepsilon.$$

Since this is true for all $\varepsilon \in (0, \bar{\varepsilon}]$, it implies equality in (36).

Case 2: Suppose $g(\bar{x}) \notin \mathbb{R}$. Then $g(\bar{x}) = -\infty$, because $g(\bar{x}) = +\infty$ implies that $\bar{x}$ is a global maximum. Since $g \not\equiv -\infty$, there is $\bar{n} \in \mathbb{N}$ such that $\{g \geq -n\} \neq \emptyset$ for all $n \geq \bar{n}$. For all such n, $\bar{x} \notin \{g \geq -n\} \neq \emptyset$. Then there exists $\bar{\xi} = \bar{\xi}_n \in \Delta$ such that $\langle \bar{\xi}, \bar{x} \rangle < \langle \bar{\xi}, x \rangle$ for all $x \in \{g \geq -n\}$. That is, $\{g \geq -n\} \subseteq \{\bar{\xi} > \langle \bar{\xi}, \bar{x} \rangle\}$ and $\{\bar{\xi} \leq \langle \bar{\xi}, \bar{x} \rangle\} \subseteq \{g < -n\}$. Thus $G_{\bar{\xi}}\left(\langle \bar{\xi}, \bar{x} \rangle\right) \leq -n$ and

$$g(\bar{x}) \leq \inf_{\xi \in \Delta} G_{\xi}\left(\langle \xi, \bar{x} \rangle\right) \leq G_{\bar{\xi}}\left(\langle \bar{\xi}, \bar{x} \rangle\right) \leq -n.$$

Since this is true for all $n \geq \bar{n}$, $\inf_{\xi \in \Delta} G_{\xi}\left(\langle \xi, \bar{x} \rangle\right) = -\infty = g(\bar{x})$, and again equality holds in (36).

"Only if". Suppose (34) holds, i.e., $g(x) = \inf_{\xi \in \Delta} G_{\xi}\left(\langle \xi, x \rangle\right)$ for all $x \in X$.

Let $r \in \mathbb{R}$ and $\bar{x} \notin \{g \geq r\}$, i.e., $g(\bar{x}) < r$. It follows that there is $\bar{\xi} \in \Delta$ for which $G_{\bar{\xi}}\left(\langle \bar{\xi}, \bar{x} \rangle\right) < r$. If there is $y \in \{g \geq r\}$ such that $\langle \bar{\xi}, y \rangle \leq \langle \bar{\xi}, \bar{x} \rangle$, then $g(y) \leq G_{\bar{\xi}}\left(\langle \bar{\xi}, \bar{x} \rangle\right) < r$, a contradiction. Therefore, $\langle \bar{\xi}, \bar{x} \rangle < \langle \bar{\xi}, y \rangle$ for all $y \in \{g \geq r\}$. By Lemma 24, $\{g \geq r\}$ is evenly convex. Since this is true for all $r \in \mathbb{R}$, g is evenly quasiconcave.

If $x \geq y$, then $\langle \xi, x \rangle \geq \langle \xi, y \rangle$ for all $\xi \in \Delta$, $G_{\xi}\left(\langle \xi, x \rangle\right) \geq G_{\xi}\left(\langle \xi, y \rangle\right)$ for all $\xi \in \Delta$, and $g(x) \geq g(y)$. That is g is monotone.

(i) Suppose that g is lower semicontinuous, then – by Lemma 29 – the map $(t, \xi) \mapsto G_{\xi}(t)$ is lower semicontinuous on $\mathbb{R} \times \Delta$. For all $x \in X$, the map $\xi \mapsto \left(\langle \xi, x \rangle, \xi\right)$ is continuous, thus their composition $\xi \mapsto G_{\xi}\left(\langle \xi, x \rangle\right)$ is lower semicontinuous and, by the Weierstrass Theorem, it admits minimum point on the compact set Δ.

(ii) Let $\bar{x} \in X$. If $\bar{x}$ is a global maximum for g on X, then, by (35) and the definition of upper semicontinuous envelope,

$$g(\bar{x}) \leq G_{\xi}\left(\langle \xi, \bar{x} \rangle\right) \leq G_{\xi}^{+}\left(\langle \xi, \bar{x} \rangle\right) \leq G_{\xi}\left(\langle \xi, \bar{x} \rangle + 1\right) \leq g(\bar{x}), \qquad \forall \xi \in \Delta,$$

and $g(\bar{x}) = \inf_{\xi \in \Delta} G_{\xi}^{+}\left(\langle \xi, \bar{x} \rangle\right)$.

If $\bar{x}$ is not a global maximum for g on X. There exists a sequence $\{r_n\} \subseteq \mathbb{R}$ such that $r_n \downarrow g(\bar{x})$ and $\bar{x} \notin \{g \geq r_n\}$ (that is $g(\bar{x}) < r_n$) for all $n \in \mathbb{N}$ (in fact, it cannot be $g(\bar{x}) = \infty$). Moreover, since there exists $\bar{y} \in X$ such that $g(\bar{y}) > g(\bar{x})$, eventually $g(\bar{y}) > r_n$ and $\{g \geq r_n\} \neq \emptyset$. Wlog $\{g \geq r_n\} \neq \emptyset$ for all $n \in \mathbb{N}$.

Let $n \in \mathbb{N}$. Since $\{g \geq r_n\}$ is closed, convex, and nonempty, by a strong separation theorem there are $\xi_n \in X^*$ and $\varepsilon_n > 0$ such that $\langle \xi_n, \bar{x} \rangle + \varepsilon_n < \langle \xi_n, x \rangle$ for all $x \in \{g \geq r_n\}$. Since $\xi_n \neq 0$, then $\langle \xi_n / \|\xi_n\|, \bar{x} \rangle + \varepsilon_n / \|\xi_n\| < \langle \xi_n / \|\xi_n\|, x \rangle$ for all $x \in \{g \geq r_n\}$. Wlog ξ_n belongs to the unit ball in X^*. Next we show that ξ_n is positive, thus (wlog) $\xi_n \in \Delta$. Let $z \in X_+$ and take $y \in \{g \geq r_n\}$. Notice that, by monotonicity, $y + mz \in \{g \geq r_n\}$ for all $m \in \mathbb{N}$, and so $\langle \xi_n, \bar{x} \rangle + \varepsilon_n < \langle \xi_n, y + mz \rangle$. Then, $\langle \xi_n, z \rangle > m^{-1}\left(\langle \xi_n, \bar{x} \rangle + \varepsilon_n - \langle \xi_n, y \rangle\right)$ for all $m \in \mathbb{N}$, which implies $\langle \xi_n, z \rangle \geq 0$. This is true for all $n \in \mathbb{N}$.

Therefore, $\{g \geq r_n\} \subseteq \{\xi_n > \langle \xi_n, \bar{x} \rangle + \varepsilon_n\}$ with $\xi_n \in \Delta$ and $\varepsilon_n > 0$ for all $n \in \mathbb{N}$. That is, $\{\xi_n \leq \langle \xi_n, \bar{x} \rangle + \varepsilon_n\} \subseteq \{g < r_n\}$. This implies $G_{\xi_n}\left(\langle \xi_n, \bar{x} \rangle + \varepsilon_n\right) \leq r_n$. Therefore, for all $n \in \mathbb{N}$,

$$g(\bar{x}) = \inf_{\xi \in \Delta} G_{\xi}\left(\langle \xi, \bar{x} \rangle\right) \leq \inf_{\xi \in \Delta} G_{\xi}^{+}\left(\langle \xi, \bar{x} \rangle\right) \leq G_{\xi_n}^{+}\left(\langle \xi_n, \bar{x} \rangle\right) \leq G_{\xi_n}\left(\langle \xi_n, \bar{x} \rangle + \varepsilon_n\right) \leq r_n,$$

which yields the result. ■

The next result considers the representation (34) for a monotone function defined on a subset Y.

Theorem 34 *Let $g : Y \to \mathbb{R}$ be a quasiconcave and monotone function defined on a convex subset Y of X. Then,*

$$g(y) = \inf_{\xi \in \Delta} G_\xi(\langle \xi, y \rangle) = \inf_{\xi \in \Delta} g_\xi(\langle \xi, y \rangle), \qquad \forall y \in Y, \tag{37}$$

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 infima in (37) are attained for all $y \in Y$.

Proof. (i) Suppose that g is lower semicontinuous and that Y is lower open. We want to prove (37) with min in place of inf. The function $\hat{g} : X \to [-\infty, \infty]$ defined by

$$\hat{g}(x) = \sup\{g(y) : Y \ni y \leq x\}. \tag{38}$$

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

Assume first that Y is open. Since

$$\{x \in X : \hat{g}(x) > t\} = \{y \in Y : g(y) > t\} + X_+, \qquad \forall t \in \mathbb{R},$$

the function $\hat{g}$ is quasiconcave and lower semicontinuous. By Theorem 33,

$$\hat{g}(x) = \min_{\xi \in \Delta} \hat{G}_\xi(\langle \xi, x \rangle) = \min_{\xi \in \Delta} \hat{g}_\xi(\langle \xi, x \rangle),$$

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, $g(y) = \min_{\xi \in \Delta} G_\xi(\langle \xi, y \rangle)$, and so the first part of (37) holds. Analogously, there is $\zeta_y \in \Delta$ such that

$$\hat{g}(y) = \hat{g}_{\zeta_y}(\langle \zeta_y, y \rangle) \geq g_{\zeta_y}(\langle \zeta_y, y \rangle) \geq g(y) = \hat{g}(y).$$

Hence, $g(y) = \min_{\xi \in \Delta} g_\xi(\langle \xi, y \rangle)$, and so the second part of (37) holds.

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 27). Moreover, $\tilde{g}$ extends g. In fact, for all $y \in Y$,

$$\tilde{g}(y) = \sup_n \hat{g}\left(y - n^{-1}e\right) = \lim_n \hat{g}\left(y - n^{-1}e\right) = \lim_n g\left(y - n^{-1}e\right)$$

since eventually $y - n^{-1}e \in Y$, and by monotonicity and lower semicontinuity of g on Y

$$g(y) \geq \lim_n g\left(y - n^{-1}e\right) \geq g(y).$$

By proceeding as in the first part of the proof, we can then prove that (37) holds.[25]

(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 (38). 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 26, its upper semicontinuous envelope $\hat{g}^+$ is monotone and quasiconcave too. Denote it by $\bar{g}$. Next we show that $\bar{g}$ extends g.

[25] By Theorem 33, $\tilde{g}(x) = \min_{\xi \in \Delta} \tilde{G}_\xi(\langle \xi, x \rangle)$, 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, $g(y) = \min_{\xi \in \Delta} G_\xi(\langle \xi, y \rangle)$, and analogously...

- If Y is upper open. Let $y \in Y$, then

$$\bar{g}(y) = \inf_n \hat{g}\left(y + n^{-1}e\right) = \lim_n \hat{g}\left(y + n^{-1}e\right) = \lim_n g\left(y + n^{-1}e\right)$$

since eventually $y + n^{-1}e \in Y$, and by monotonicity and lower semicontinuity of g on Y

$$g(y) \leq \lim_n g\left(y + n^{-1}e\right) \leq g(y).$$

- 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

$$\hat{g}(x) = g(x \wedge z), \qquad \forall x \in X_+ + w.$$

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.

For all $\xi \in \Delta$ and $t \in \mathbb{R}$,

$$\begin{aligned} G_\xi(t) &= \sup\{g(x) : x \in Y, \langle \xi, x\rangle \leq t\} = \sup\{\bar{g}(x) : x \in Y, \langle \xi, x\rangle \leq t\} \\ &\leq \sup\{\bar{g}(x) : x \in X, \langle \xi, x\rangle \leq t\} = \bar{G}_\xi(t). \end{aligned}$$

By Theorem 33, for all $y \in Y$,

$$g(y) = \bar{g}(y) = \inf_{\xi \in \Delta} \bar{G}_\xi(\langle \xi, y\rangle) \geq \inf_{\xi \in \Delta} G_\xi(\langle \xi, y\rangle) \geq \inf_{\xi \in \Delta} g_\xi(\langle \xi, y\rangle) \geq g(y),$$

as desired. ■

Corollary 35 *Let $g : X(T) \to \mathbb{R}$ be quasiconcave and monotone. If g is continuous, then*

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

In particular, the infima are attained if T is lower open.

A.2.2 Concavity

The next two corollaries of Theorem 33, 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 36 *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_ξ is concave for each $\xi \in \Delta$;*

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

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

Proof.* First notice that, for all $(\lambda, \xi) \in \mathbb{R} \times \Delta$,

$$
\begin{aligned}
g^*(\lambda\xi) &= \inf_{x \in X} \{\lambda \langle \xi, x\rangle - g(x)\} = \inf_{t \in \mathbb{R}} \inf_{\{x \in X : \langle \xi, x\rangle = t\}} \{\lambda \langle \xi, x\rangle - g(x)\} \\
&= \inf_{t \in \mathbb{R}} \inf_{\{x \in X : \langle \xi, x\rangle = t\}} \{\lambda t - g(x)\} = \inf_{t \in \mathbb{R}} \left\{ \lambda t - \sup_{\{x \in X : \langle \xi, x\rangle = t\}} g(x) \right\} \\
&= \inf_{t \in \mathbb{R}} \{\lambda t - g_\xi(t)\} = (g_\xi)^*(\lambda).
\end{aligned}
$$

Moreover, (ii) is equivalent to (iii) since $G_\xi = g_\xi$ for all $\xi \in \Delta$, by monotonicity of g.

(i) implies (iii). In fact, for all $t, r \in \mathbb{R}$ and $\alpha \in (0, 1)$,

$$
\begin{aligned}
G_\xi(\alpha t + (1-\alpha) r) &= \sup \{g(x) : x \in X,\ \langle \xi, x\rangle \le \alpha t + (1-\alpha) r\} \\
&\ge \sup \{g(\alpha y + (1-\alpha) z) : y, z \in X,\ \langle \xi, y\rangle \le t,\ \langle \xi, z\rangle \le r\} \\
&\ge \sup \{\alpha g(y) + (1-\alpha) g(z) : y, z \in X,\ \langle \xi, y\rangle \le t,\ \langle \xi, z\rangle \le r\} \\
&= \alpha \sup \{g(y) : y \in X,\ \langle \xi, y\rangle \le t\} + (1-\alpha) \sup \{g(z) : z \in X,\ \langle \xi, z\rangle \le r\} \\
&= \alpha G_\xi(t) + (1-\alpha) G_\xi(r).
\end{aligned}
$$

(ii) implies (iv). Let $\xi \in \Delta$. Since $g_\xi : \mathbb{R} \to (-\infty, \infty]$ and it is concave, then either g_ξ is finite on $\mathbb{R}$ (hence continuous), or $g_\xi \equiv \infty$. In either cases

$$
g_\xi(t) = \inf_{\lambda \in \mathbb{R}} \left\{\lambda t - (g_\xi)^*(\lambda)\right\}
$$

for all $t \in \mathbb{R}$.[26] Monotonicity of g_ξ implies $(g_\xi)^*(\lambda) = -\infty$ for all $\lambda < 0$, whence

$$
g_\xi(t) = \inf_{\lambda \ge 0} \left\{\lambda t - (g_\xi)^*(\lambda)\right\} = \inf_{\lambda \ge 0} \{\lambda t - g^*(\lambda \xi)\}
$$

for all $t \in \mathbb{R}$.

(iv) implies (i). Assume $g_\xi(t) = \inf_{\lambda \in \mathbb{R}_+} (\lambda t - g^*(\lambda\xi))$ for each $(t, \xi) \in \mathbb{R} \times \Delta$. By Theorem 33,

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

By (iv) g_ξ is concave, for all $\xi \in \Delta$. Therefore

$$
x \mapsto g_\xi(\langle \xi, x\rangle), \qquad \forall x \in X,
$$

is concave for all $\xi \in \Delta$, and g (being an infimum of concave functions) is concave too. ■

Next we consider normalized functions.

Corollary 37 *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:*

[26] If $g_\xi \equiv \infty$, then $g_\xi^*(\lambda) = \inf_{t \in \mathbb{R}} \{\lambda t - \infty\} = -\infty$ for all $\lambda \in \mathbb{R}$ and $\inf_{\lambda \in \mathbb{R}} (\lambda t - (-\infty)) = \infty$ for all $t \in \mathbb{R}$.

(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* $\xi \in \Delta$.

Proof.* By Theorem 33, if g is normalized,

$$t = g(te) = \inf_{\xi\in\Delta} G_\xi(\langle \xi, te\rangle) = \inf_{\xi\in\Delta} G_\xi(t), \qquad \forall t \in \mathbb{R}.$$

Conversely, if $\inf_{\xi\in\Delta} G_\xi(t) = t$ for all $t \in \mathbb{R}$, then

$$g(te) = \inf_{\xi\in\Delta} G_\xi(\langle \xi, te\rangle) = \inf_{\xi\in\Delta} G_\xi(t) = t, \qquad \forall t \in \mathbb{R},$$

as desired.

That (ii) implies (i) is well known.

(i) implies (iii). Let $\xi \in \Delta$. Since g is normalized and monotone, it is real-valued. By Corollary 36,

$$g_\xi(t) = \inf_{\lambda\in\mathbb{R}_+} \{\lambda t - g^*(\lambda\xi)\}, \qquad \forall (t,\xi) \in \mathbb{R}\times\Delta. \tag{39}$$

Since g is monotone and normalized, $g^*(\xi) = -\infty$ if $\xi \notin \Delta$. Hence, $g^*(\lambda\xi) = -\infty$ if $\xi \in \Delta$ and $\lambda \neq 1$, and so, by (39), $g_\xi(t) = t - g^*(\xi)$.

(iii) implies (ii). By Theorem 33 we have:

$$g(x) = \inf_{\xi\in\Delta} (\langle \xi, x\rangle - g^*(\xi)), \qquad \forall x \in X.$$

Which clearly implies translation invariance. ■

A.2.3 Topological Representation

Next we give a topological version of Theorem 33. Also in this case we only consider the monotone case, and we refer to [7] for the general case and for a proof.

Theorem 38 *A function* $g : X \to \mathbb{R}$ *is uniformly continuous, quasiconcave, and monotone if and only if*

$$g(x) = \min_{\xi\in\Delta} G_\xi(\langle \xi, x\rangle) = \min_{\xi\in\Delta} g_\xi(\langle \xi, x\rangle), \qquad \forall x \in X, \tag{40}$$

$\mathrm{dom}(G_\xi) \in \{\emptyset, \mathbb{R}\}$ *for all* $\xi \in \Delta$, *and* $\{G_\xi\}_{\xi\in\Delta:\mathrm{dom}(G_\xi)=\mathbb{R}}$ *are uniformly equicontinuous.*[27]

Proof.* Suppose g is quasiconcave and uniformly continuous. In particular, g is lower semicontinuous and, by Theorem 33, we have the representation (40). As g is uniformly continuous, for all $\varepsilon > 0$, there is some $\delta > 0$ such that $\|x - y\| \le \delta$ implies $|g(x) - g(y)| \le \varepsilon$. In particular,

$$g(x + \delta e) \le g(x) + \varepsilon \text{ and } g(x - \delta e) \ge g(x) - \varepsilon. \tag{41}$$

Let $\xi \in \Delta$, $t \in \mathrm{dom}(G_\xi)$, and $t' \in \mathbb{R}$ with $|t - t'| \le \delta$. Consider two cases:

Case 1: $t' \le t$. Then,

$$\begin{aligned} G_\xi(t) - \varepsilon &= \sup\{g(x) - \varepsilon : x \in X, \langle \xi, x\rangle \le t\} \le \sup\{g(x - \delta e) : x \in X, \langle \xi, x\rangle \le t\} \\ &= \sup\{g(y) : y \in X, \langle \xi, y + \delta e\rangle \le t\} = \sup\{g(x) : x \in X, \langle \xi, x\rangle \le t - \delta\} \\ &= G_\xi(t - \delta) \le G_\xi(t') \le G_\xi(t). \end{aligned}$$

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

Therefore, $|G_\xi(t) - G_\xi(t')| \leq \varepsilon$.

Case 2: $t' \geq t$. Then,

$$\begin{aligned} G_\xi(t) \leq G_\xi(t') \leq G_\xi(t+\delta) &= \sup\{g(x) : x \in X, \langle \xi, x\rangle \leq t+\delta\} \\ &= \sup\{g(x) : x \in X, \langle \xi, x-\delta e\rangle \leq t\} = \sup\{g(y+\delta e) : y \in X, \langle \xi, y\rangle \leq t\} \\ &\leq \sup\{g(y)+\varepsilon : y \in X, \langle \xi, y\rangle \leq t\} = G_\xi(t)+\varepsilon, \end{aligned}$$

and again $|G_\xi(t) - G_\xi(t')| \leq \varepsilon$.

Therefore:

- If $\operatorname{dom}(G_\xi) \neq \emptyset$, for all $t \in \operatorname{dom}(G_\xi)$, then $[t-\delta, t+\delta] \subseteq \operatorname{dom}(G_\xi)$; that is, $\operatorname{dom}(G_\xi) = \mathbb{R}$.
- For all $\xi \in \Delta$ with $\operatorname{dom}(G_\xi) = \mathbb{R}$, $|t-t'| \leq \delta$ implies $|G_\xi(t) - G_\xi(t')| \leq \varepsilon$.

As wanted.

As to the converse, assume that (40) holds, $\operatorname{dom}(G_\xi) \in \{\emptyset, \mathbb{R}\}$ for all $\xi \in \Delta$, and $\{G_\xi\}_{\xi\in\Delta:\operatorname{dom}(G_\xi)=\mathbb{R}}$ are uniformly equicontinuous. By Theorem 33, g is evenly quasiconcave and monotone (while by assumption it is real-valued). Moreover, for all $\varepsilon > 0$, there is $\delta > 0$ such that $|G_\xi(t) - G_\xi(t')| \leq \varepsilon$ for all $t, t' \in \mathbb{R}$ with $|t-t'| \leq \delta$ and all $\xi \in \Delta$ with $\operatorname{dom}(G_\xi) = \mathbb{R}$. Take $x, y \in X$ such that $\|x-y\| \leq \delta$. There is $\xi_x \in \Delta$ such that $g(x) = G_{\xi_x}(\langle \xi_x, x\rangle)$. Since $g(x) \in \mathbb{R}$, then $\operatorname{dom}(G_{\xi_x}) = \mathbb{R}$. Moreover, if $\|x-y\| \leq \delta$, then $|\langle \xi_x, x\rangle - \langle \xi_x, y\rangle| \leq \|\xi_x\| \|x-y\| \leq \delta$. By uniform equicontinuity, $|G_{\xi_x}(\langle \xi_x, x\rangle) - G_{\xi_x}(\langle \xi_x, y\rangle)| \leq \varepsilon$, and so

$$g(x) = G_{\xi_x}(\langle \xi_x, x\rangle) \geq G_{\xi_x}(\langle \xi_x, y\rangle) - \varepsilon \geq \min_{\xi\in\Delta} G_\xi(\langle \xi, y\rangle) - \varepsilon = g(y) - \varepsilon.$$

Exchanging the two points x and y, we obtain $|g(x) - g(y)| \leq \varepsilon$, and so g is uniformly continuous.■

A.2.4 Uniqueness

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

Lemma 39 *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,

$$G(t, \xi) = \sup_{x\in X:\langle \xi, x\rangle \leq t} g(x) = G_\xi(t), \qquad \forall (t, \xi) \in \mathbb{R} \times \Delta. \tag{42}$$

Proof.* Observe first that there is no loss of generality in assuming that g and G are real-valued (bounded) functions. For, let $\varphi : [-\infty, \infty] \to \mathbb{R}$ be a strictly increasing, extended-valued continuous, and bounded function, say $\varphi(t) = \arctan t$. If we consider $h = \varphi \circ g$ and $H = \varphi \circ G$, they satisfy (i)-(iv) and are real-valued (bounded), and $H(t, \xi) = \sup_{x\in X:\langle \xi, x\rangle \leq t} h(x)$ implies (42), by Lemma 32.

Fix $\bar{\xi} \in \Delta$ and $t \in \mathbb{R}$. We have

$$G_{\bar{\xi}}(t) = \sup_{x \in X: \langle \bar{\xi}, x \rangle \leq t} g(x) = \sup_{x \in X: \langle \bar{\xi}, x \rangle \leq t} \inf_{\xi \in \Delta} G(\langle \xi, x \rangle, \xi).$$

Define $\Gamma : X \times \Delta \to \mathbb{R}$ by $\Gamma(x, \xi) = G(\langle \xi, x \rangle, \xi)$ for all $(x, \xi) \in X \times \Delta$. It is easy to show that Γ is lower semicontinuous on $X \times \Delta$ (see, e.g., Claim 1 of Lemma 51). Moreover, $\Gamma(x, \cdot) : \Delta \to \mathbb{R}$ is quasiconvex on Δ for all $x \in X$, and, by assumption (ii), $\Gamma(\cdot, \xi) : X \to \mathbb{R}$ is quasiconcave on X for all $\xi \in \Delta$. As $Y = \left\{x \in X : \langle \bar{\xi}, x \rangle \leq t\right\}$ is nonempty and convex, and Δ is nonempty, convex, and compact, we can invoke a well known Minimax Theorem (e.g., [25, Theorem 4]) and obtain

$$G_{\bar{\xi}}(t) = \sup_{x \in X: \langle \bar{\xi}, x \rangle \leq t} \inf_{\xi \in \Delta} G(\langle \xi, x \rangle, \xi) = \sup_{x \in Y} \inf_{\xi \in \Delta} \Gamma(x, \xi) = \inf_{\xi \in \Delta} \sup_{x \in Y} \Gamma(x, \xi) = \inf_{\xi \in \Delta} \sup_{x \in X: \langle \bar{\xi}, x \rangle \leq t} G(\langle \xi, x \rangle, \xi).$$

It remains to study the program

$$\sup_{x \in X: \langle \bar{\xi}, x \rangle \leq t} G(\langle \xi, x \rangle, \xi).$$

Consider separately two cases:

(a) $\xi \in \langle \bar{\xi} \rangle$. There exists $\lambda \in \mathbb{R}$ such that $\xi = \lambda \bar{\xi}$, but $1 = \xi(e) = \lambda \bar{\xi}(e) = \lambda$, implies $\xi = \bar{\xi}$. Thus,

$$\sup_{x \in X: \langle \bar{\xi}, x \rangle \leq t} G(\langle \xi, x \rangle, \xi) = \sup_{x \in X: \langle \bar{\xi}, x \rangle \leq t} G\left(\langle \bar{\xi}, x \rangle, \bar{\xi}\right) = G\left(t, \bar{\xi}\right),$$

attained, for example, at $\bar{x} = te$.

(b) $\xi \notin \langle \bar{\xi} \rangle$. Hence, $\ker\left(\bar{\xi}\right) \nsubseteq \ker(\xi)$, i.e., there is $y \in X$ for which $\langle \bar{\xi}, y \rangle = 0$ and $\langle \xi, y \rangle \neq 0$. This implies that, choosing $\bar{x}$ such that $\langle \bar{\xi}, \bar{x} \rangle = t$, the straight line $\{\bar{x} + \alpha y : \alpha \in \mathbb{R}\}$ is included into the hyperplane $\langle \bar{\xi}, x \rangle = t$. Thus,

$$\sup_{r \in \mathbb{R}} G(r, \xi) \geq \sup_{x \in X: \langle \bar{\xi}, x \rangle \leq t} G(\langle \xi, x \rangle, \xi) \geq \sup_{\alpha \in \mathbb{R}} G(\langle \xi, \bar{x} + \alpha y \rangle, \xi) = \sup_{r \in \mathbb{R}} G(r, \xi)$$

hence, in view of (i) and (ii),

$$\sup_{x \in X: \langle \bar{\xi}, x \rangle \leq t} G(\langle \xi, x \rangle, \xi) = \sup_{r \in \mathbb{R}} G(r, \xi) = \lim_{r \to \infty} G(r, \xi) = \lim_{r \to \infty} G\left(r, \bar{\xi}\right) \geq G\left(t, \bar{\xi}\right).$$

Conclude that

$$G_{\bar{\xi}}(t) = \inf_{\xi \in \Delta} \left(\sup_{x \in X: \langle \bar{\xi}, x \rangle \leq t} G(\langle \xi, x \rangle, \xi) \right) = G\left(t, \bar{\xi}\right),$$

as desired. ■

Remark 40 *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 38 takes the following form for positively homogeneous and quasiconcave functionals such that $g(x) \neq 0$ for some $x \in X_+$.

Theorem 41 *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*

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

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

$$G_\xi(t) = \begin{cases} \frac{t}{c_1(\xi)} & \text{if } t \geq 0 \text{ and } \xi \in \widetilde{\Delta} \\ \frac{t}{c_2(\xi)} & \text{if } t \leq 0 \text{ and } \xi \in \widetilde{\Delta} \\ \infty & \text{if } \xi \in \Delta \setminus \widetilde{\Delta}. \end{cases} \tag{44}$$

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

Proof.* We build on some Lemmas.

Lemma 42 *Let $(a_i)_{i \in I}, (b_i)_{i \in I} \subseteq \mathbb{R}^+$. The family of functions*

$$f_i(t) = \begin{cases} a_i t & \text{if } t \geq 0 \\ b_i t & \text{if } t \leq 0 \end{cases}$$

is uniformly equicontinuous if and only if $\sup_{i \in I} a_i, \sup_{i \in I} b_i < \infty$.

Proof of Lemma 42.* First observe that a family of monotone functions is uniformly equicontinuous if and only if for every $\varepsilon > 0$ there is $\delta > 0$ such that

$$f_i(t + \delta) \leq f_i(t) + \varepsilon \tag{45}$$

for all $t \in \mathbb{R}$ and $i \in I$.[28]

In our special case for all $i \in I$, $t \in \mathbb{R}$, and $\delta > 0$,

$$\begin{aligned} f_i(t+\delta) - f_i(t) &= \begin{cases} a_i \delta & \text{if } t \geq 0 \\ a_i t + a_i \delta - b_i t & \text{if } -\delta < t < 0 \\ b_i \delta & \text{if } t + \delta \leq 0 \text{ (i.e. } t \leq -\delta) \end{cases} \\ (a_i t < 0 \text{ if } -\delta < t < 0) &\leq \begin{cases} a_i \delta & \text{if } t \geq 0 \\ a_i \delta - b_i t & \text{if } -\delta < t < 0 \\ b_i \delta & \text{if } t \leq -\delta \end{cases} \\ (b_i \delta \geq -b_i t \geq 0 \text{ if } -\delta < t < 0) &\leq \begin{cases} a_i \delta & \text{if } t \geq 0 \\ a_i \delta + b_i \delta & \text{if } -\delta < t < 0 \\ b_i \delta & \text{if } t \leq -\delta \end{cases} \\ &\leq (a_i + b_i)\delta \leq \left(\sup_{i \in I} a_i + \sup_{i \in I} b_i + 1 \right) \delta. \end{aligned}$$

Therefore, if $\sup_{i \in I} a_i, \sup_{i \in I} b_i < \infty$, for all $\varepsilon > 0$ it suffices to take

$$\delta < \frac{\varepsilon}{(\sup_{i \in I} a_i + \sup_{i \in I} b_i + 1)}$$

[28] If for every $\varepsilon > 0$ there is $\delta > 0$ such that $|t' - t''| \leq \delta$ implies $|f_i(t') - f_i(t'')| \leq \varepsilon$, for all $t', t'' \in \mathbb{R}$ and all $i \in I$, then $f_i(t+\delta) - f_i(t) = |f_i(t+\delta) - f_i(t)| \leq \varepsilon$ for all $t \in \mathbb{R}$ and all $i \in I$. Conversely, if condition (45) holds, consider $t', t'' \in \mathbb{R}$ with $|t' - t''| \leq \delta$, wlog $t' \geq t''$, then $t' \leq t'' + \delta$, $f_i(t'' + \delta) \leq f_i(t'') + \varepsilon$, and monotonicity, deliver $f_i(t') \leq f_i(t'' + \delta) \leq f_i(t'') + \varepsilon$, whence $|f_i(t') - f_i(t'')| = f_i(t') - f_i(t'') \leq \varepsilon$ for all $i \in I$.

to obtain

$$f_i(t+\delta) - f_i(t) \le \left(\sup_{i\in I} a_i + \sup_{i\in I} b_i + 1\right)\delta \le \varepsilon$$

for all $i \in I$, $t \in \mathbb{R}$, which implies uniform equicontinuity.

If $\sup_{i\in I} a_i = \infty$, then for all $\delta > 0$

$$f_i(0+\delta) - f_i(0) = a_i\delta$$

and hence $\sup_{i\in I}(f_i(0+\delta) - f_i(0)) = \infty$, which contradicts condition (45).

If $\sup_{i\in I} b_i = \infty$, then for all $\delta > 0$ take $t_\delta < -\delta$

$$f_i(t_\delta+\delta) - f_i(t_\delta) = b_i\delta$$

and hence $\sup_{i\in I}(f_i(t_\delta+\delta) - f_i(t_\delta)) = \infty$, which contradicts condition (45). □

Lemma 43 *Let C be a convex subset of a vector space and $f_1, f_2 : C \to \mathbb{R}$ be quasiconvex functions. If $f_1 \ge 0$, $f_2 \le 0$, and $f_1 f_2 = 0$, then $f_1 + f_2$ is quasiconvex.*

Proof of Lemma 43.* Let $f = f_1 + f_2$. Set $C^- = \{x \in C : f_2(x) < 0\}$. The set C^- is convex, and we can assume $C^- \neq \emptyset$ (else $f_1 + f_2 = f_1$ is quasiconvex). As f_1 and f_2 are quasiconvex, we have

$$\begin{aligned} f_1(\alpha x_1 + (1-\alpha)x_2) &\le f_1(x_1) \vee f_1(x_2) \text{ and} \\ f_2(\alpha x_1 + (1-\alpha)x_2) &\le f_2(x_1) \vee f_2(x_2)\text{, then} \\ f(\alpha x_1 + (1-\alpha)x_2) &\le f_1(x_1) \vee f_1(x_2) + f_2(x_1) \vee f_2(x_2) \end{aligned}$$

for all $x_1, x_2 \in C$ and $\alpha \in [0,1]$; we want to show that $f(\alpha x_1 + (1-\alpha)x_2) \le f(x_1) \vee f(x_2)$.

Consider the following cases:

(a) $x_1, x_2 \in C^-$;

(b) $x_1 \in C^-$ and $x_2 \notin C^-$;

(c) $x_1, x_2 \notin C^-$.

Case (a). The convexity of C^- implies $\alpha x_1 + (1-\alpha)x_2 \in C^-$, and over C^- we have $f_1 = 0$, then $f_{|C^-} = f_2$ delivers the result.

Case (b), we have $f_1(x_1) = 0$, $f_2(x_1) < 0$, $f_2(x_2) = 0$, and $f_1(x_2) \ge 0$. Therefore,

$$f(\alpha x_1 + (1-\alpha)x_2) \le f_1(x_1) \vee f_1(x_2) + f_2(x_1) \vee f_2(x_2) = f_1(x_2)$$

and

$$f(x_1) \vee f(x_2) = (0 + f_2(x_1)) \vee (f_1(x_2) + 0) = f_1(x_2)$$

as wanted.

Case (c) we have $f_2(x_1) = f_2(x_2) = 0$, $f_1(x_1) \ge 0$, $f_1(x_2) \ge 0$. Hence,

$$f(\alpha x_1 + (1-\alpha)x_2) \le f_1(x_1) \vee f_1(x_2) + f_2(x_1) \vee f_2(x_2) = f_1(x_1) \vee f_1(x_2) = f(x_1) \vee f(x_2)$$

which concludes the proof. □

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*

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

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 $\xi \in \widetilde{\Delta}$.*

Proof of Lemma 44.* Next we show that G satisfies all the conditions of Lemma 39.

(i) If $\xi \in \widetilde{\Delta}$, then

$$\lim_{t\to\infty} G(t,\xi) = \lim_{t\to\infty} \frac{t}{c_1(\xi)} = \infty$$

since $c_1(\xi) \in (0,\infty)$, and the same is true if $\xi \notin \widetilde{\Delta}$.

(ii) $G(\cdot,\xi)$ is obviously monotone and extended-valued continuous on $\mathbb{R}$, for each $\xi \in \Delta$.

(iii) Next we check that $G(t,\xi)$ is lower semicontinuous and quasiconvex on $\mathbb{R} \times \Delta$. Notice that

$$G(t,\xi) = \frac{t^+}{c_1(\xi)} - \frac{t^-}{c_2(\xi)}, \tag{49}$$

for all $(t,\xi) \in \mathbb{R} \times \widetilde{\Delta}$. Since $\mathbb{R} \times \widetilde{\Delta}$ is closed and convex in $\mathbb{R} \times \Delta$, and $G(t,\xi) = \infty$ outside $\mathbb{R} \times \widetilde{\Delta}$, it suffices to check that G is lower semicontinuous and quasiconvex on $\mathbb{R} \times \widetilde{\Delta}$, where G is finite and given by (49). It is convenient to study first separately the two functions $(t,\xi) \mapsto t^+/c_1(\xi)$, and $(t,\xi) \mapsto t^-/c_2(\xi)$.

For all $\alpha \geq 0$, we have

$$\left\{(t,\xi) \in \mathbb{R} \times \widetilde{\Delta} : \frac{t^+}{c_1(\xi)} \leq \alpha\right\} = \left\{(t,\xi) \in \mathbb{R} \times \widetilde{\Delta} : t^+ - \alpha c_1(\xi) \leq 0\right\}$$

and the latter set is closed and convex since the functions $t \mapsto t^+$ and $\xi \mapsto -\alpha c_1(\xi)$ are convex and lower semicontinuous; while for $\alpha < 0$ the set is empty. Therefore $(t,\xi) \mapsto t^+/c_1(\xi)$ is lower semicontinuous and quasiconvex.

Analogously, for all $\alpha > 0$, we have

$$\begin{aligned}
&\left\{(t,\xi) \in \mathbb{R} \times \widetilde{\Delta} : \frac{t^-}{c_2(\xi)} \geq \alpha\right\} = \\
&= \left\{(t,\xi) \in \mathbb{R} \times \widetilde{\Delta} : \alpha c_2(\xi) - t^- \leq 0\right\} \\
&= \left\{(t,\xi) \in (-\infty,0) \times \widetilde{\Delta} : \alpha c_2(\xi) + t \leq 0\right\} \cup \left\{(t,\xi) \in \mathbb{R}^+ \times \widetilde{\Delta} : \alpha c_2(\xi) \leq 0\right\} \\
&= \left\{(t,\xi) \in (-\infty,0) \times \widetilde{\Delta} : \alpha c_2(\xi) + t \leq 0\right\} \cup \emptyset \\
&= \left\{(t,\xi) \in (-\infty,0) \times \widetilde{\Delta} : \alpha c_2(\xi) + t \leq 0\right\} \cup \left\{(t,\xi) \in \mathbb{R}^+ \times \widetilde{\Delta} : \alpha c_2(\xi) + t \leq 0\right\} \\
&= \left\{(t,\xi) \in \mathbb{R} \times \widetilde{\Delta} : \alpha c_2(\xi) + t \leq 0\right\}
\end{aligned}$$

and the latter set is closed and convex since the functions $t \mapsto t$ and $\xi \mapsto \alpha c_2(\xi)$ are convex and lower semicontinuous; while for $\alpha \leq 0$ the set is $\mathbb{R} \times \widetilde{\Delta}$. Therefore $(t,\xi) \mapsto t^-/c_2(\xi)$ is upper semicontinuous and quasiconcave.

As a consequence, the mapping $(\xi,t) \mapsto t^+/c_1(\xi) - t^-/c_2(\xi)$ is lower semicontinuous, and quasiconvex by Lemma 43, and $(\xi,t) \mapsto G(t,\xi)$ is lower semicontinuous and quasiconvex on $\mathbb{R} \times \Delta$.

For every $x \in X$, $\xi \mapsto (\langle \xi, x\rangle, \xi)$ is affine and continuous, therefore $\xi \mapsto G(\langle \xi, x\rangle, \xi)$ is lower semicontinuous and quasiconvex on Δ.

(iv) is just (46).

Lemma 39 yields the following consequences:

- $G(t,\xi) = \sup_{x \in X : \langle \xi, x\rangle \leq t} g(x) = G_\xi(t)$ for all $(t,\xi) \in \mathbb{R} \times \Delta$.
- g is monotone, quasiconcave, and upper semicontinuous;
- since, by (iii), the inf is attained in (46), then g is finite, in particular

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

  and g is positively homogeneous.

- 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

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

  set

$$H(t,\xi) = \begin{cases} \frac{t}{d_1(\xi)} & \text{if } t \geq 0 \text{ and } \xi \in \widetilde{\Gamma} \\ \frac{t}{d_2(\xi)} & \text{if } t \leq 0 \text{ and } \xi \in \widetilde{\Gamma} \\ \infty & \text{if } \xi \in \Delta \setminus \widetilde{\Gamma} \end{cases} = \begin{cases} \frac{t^+}{d_1(\xi)} - \frac{t^-}{d_2(\xi)} & \text{if } (t,\xi) \in \mathbb{R} \times \widetilde{\Gamma} \\ \infty & \text{if } (t,\xi) \in \mathbb{R} \times \left(\Delta \setminus \widetilde{\Gamma}\right) \end{cases}.$$

  Apply the previous points and obtain $H(t,\xi) = \sup_{x \in X : \langle \xi, x\rangle \leq t} g(x) = G_\xi(t) = G(t,\xi)$ for all $(t,\xi) \in \mathbb{R} \times \Delta$. In particular,

$$\begin{gathered}
\widetilde{\Gamma} = \{\xi \in \Delta : H(0,\xi) = 0\} = \{\xi \in \Delta : G(0,\xi) = 0\} = \widetilde{\Delta}, \\
1/d_1(\xi) = H(1,\xi) = G(1,\xi) = 1/c_1(\xi), \qquad \forall \xi \in \widetilde{\Delta} \text{ i.e. } d_1 = c_1, \\
-1/d_2(\xi) = H(-1,\xi) = G(-1,\xi) = -1/c_2(\xi), \qquad \forall \xi \in \widetilde{\Delta} \text{ i.e. } d_2 = c_2.
\end{gathered}$$

- For all $x \in X$ such that $\langle \xi, x \rangle \geq 0$ for all $\xi \in \widetilde{\Delta}$,

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

  which is clearly concave and non-negative.

- g is concave if and only if G_ξ is concave for each $\xi \in \Delta$. This is automatically true, if $\xi \in \Delta \setminus \widetilde{\Delta}$, while if $\xi \in \widetilde{\Delta}$ this amounts to say that the function

$$G(t, \xi) = \begin{cases} \frac{t}{c_1(\xi)} & \text{if } t \geq 0 \\ \frac{t}{c_2(\xi)} & \text{if } t \leq 0 \end{cases}$$

  is concave, or equivalently

$$\frac{1}{c_1(\xi)} \leq \frac{1}{c_2(\xi)}$$

  thus $c_2(\xi) < \infty$ and $c_1(\xi) \geq c_2(\xi)$.

As wanted. $\square$

Let $g : X \to \mathbb{R}$ be such that $g(x) \neq 0$ for some $x \in X_+$ and assume g is monotone, quasiconcave, uniformly continuous, and positively homogeneous. Theorem 38 guarantees that (43) holds. Next we prove that $G_\xi(t)$ has the representation (44).

By Theorem 38, $\mathrm{dom}(G_\xi) \in \{\emptyset, \mathbb{R}\}$ for all $\xi \in \Delta$, and $\{G_\xi\}_{\xi \in \Delta : \mathrm{dom}(G_\xi) = \mathbb{R}}$ are uniformly equicontinuous. Set $\widetilde{\Delta} = \{\xi \in \Delta : \mathrm{dom}(G_\xi) = \mathbb{R}\}$, which is not empty since g is finite. Clearly, $G_\xi(t) = \infty$ if $\xi \in \Delta \setminus \widetilde{\Delta}$. For all $\xi \in \widetilde{\Delta}$, the functions $t \mapsto G_\xi(t)$ are (monotone and) positively homogeneous, in fact, for all $t \in \mathbb{R}$ and $\lambda > 0$,

$$G_\xi(\lambda t) = \sup_{x \in X : \langle \xi, x \rangle \leq \lambda t} g(x) = \sup_{y \in X : \langle \xi, y \rangle \leq t} g(\lambda y) = \sup_{y \in X : \langle \xi, y \rangle \leq t} \lambda g(y) = \lambda G_\xi(t)$$

since $\{x \in X : \langle \xi, x \rangle \leq \lambda t\} = \lambda \{y \in X : \langle \xi, y \rangle \leq t\}$. Therefore, there are two functions $\rho_1, \rho_2 : \widetilde{\Delta} \to [0, \infty)$ such that

$$G_\xi(t) = \begin{cases} \rho_1(\xi) t & \text{if } t \geq 0 \\ \rho_2(\xi) t & \text{if } t \leq 0 \end{cases} \tag{50}$$

if $\xi \in \widetilde{\Delta}$.

If $\rho_1(\bar{\xi}) = 0$ for some $\bar{\xi} \in \widetilde{\Delta}$, then $G_{\bar{\xi}}(t) = 0$ for all $t \geq 0$ and

$$\sup_{x \in X} g(x) = \lim_{t \to \infty} G_{\bar{\xi}}(t) = 0$$

which – together with monotonicity – contradicts the assumption $g(x) \neq 0$ for some $x \in X_+$. Thus $\rho_1 : \widetilde{\Delta} \to (0, \infty)$.

Step 1. $\widetilde{\Delta}$ is convex. Notice that $\widetilde{\Delta} = \{\xi \in \Delta : G_\xi(0) < \infty\}$. If $G_{\xi_1}(0), G_{\xi_2}(0) < \infty$, quasiconvexity of $G_\xi(t)$ implies $G_{\alpha\xi_1 + (1-\alpha)\xi_2}(0) < \infty$ for all $\alpha \in (0, 1)$ too.

Step 2. The function $c_1 : \widetilde{\Delta} \to (0, \infty)$ defined by $c_1(\xi) = 1/\rho_1(\xi)$ is concave over $\widetilde{\Delta}$ (remember that $\rho_1(\xi) > 0$). Let $\xi_1, \xi_2 \in \widetilde{\Delta}$ and $\alpha \in (0, 1)$. Choose $k_1, k_2 > 0$ such that $k_1 \rho_1(\xi_1) = k_2 \rho_1(\xi_2)$. As $G_\xi(t)$ is quasiconvex,

$$G_{\alpha\xi_1 + (1-\alpha)\xi_2}(\alpha k_1 + (1-\alpha) k_2) \leq \max\{G_{\xi_1}(k_1), G_{\xi_2}(k_2)\}. \tag{51}$$

In view of (50), this amounts to

$$(\alpha k_1 + (1-\alpha) k_2) \rho_1 (\alpha \xi_1 + (1-\alpha) \xi_2) \leq \max \{k_1 \rho_1 (\xi_1), k_2 \rho (\xi_2)\} = k_1 \rho (\xi_1),$$

and so, since $k_2/k_1 = \rho_1 (\xi_1) / \rho_1 (\xi_2)$,

$$\frac{1}{\rho_1 (\alpha \xi_1 + (1-\alpha) \xi_2)} \geq \frac{\alpha k_1 + (1-\alpha) k_2}{k_1 \rho_1 (\xi_1)} = \frac{\alpha}{\rho_1 (\xi_1)} + \frac{1-\alpha}{\rho_1 (\xi_2)}.$$

This shows that $1/\rho_1 (\xi)$ is concave. Consequently, in (50) we can write $\rho_1 (\xi) t = t/c_1 (\xi)$, where $c_1 (\xi) = 1/\rho_1 (\xi)$ is concave on $\widetilde{\Delta}$.

Step 3. The region $A = \left\{\xi \in \widetilde{\Delta} : \rho_2 (\xi) > 0\right\}$ is convex. In fact, for all $\xi \in \widetilde{\Delta}$, $\rho_2 (\xi) = -G_\xi (-1)$ and thus

$$A = \left\{\xi \in \widetilde{\Delta} : G_\xi (-1) < 0\right\}$$

is convex by quasiconvexity of $G_\xi (t)$.

Step 4. The function $c_2 : A \to (0, \infty)$ defined by $c_2 (\xi) = 1/\rho_2 (\xi)$ is convex on the set A defined above. Let $\xi_1, \xi_2 \in A \subseteq \widetilde{\Delta}$ and $\alpha \in (0,1)$. Pick $k_1, k_2 < 0$ such that $k_1 \rho_2 (\xi_1) = k_2 \rho_2 (\xi_2)$. From the quasiconvexity of $G_\xi (t)$ we have (51). Namely,

$$(\alpha k_1 + (1-\alpha) k_2) \rho_2 (\alpha \xi_1 + (1-\alpha) \xi_2) \leq \max \{k_1 \rho_2 (\xi_1), k_2 \rho_2 (\xi_2)\} = k_1 \rho_2 (\xi_1),$$

that implies

$$\frac{1}{\rho_2 (\alpha \xi_1 + (1-\alpha) \xi_2)} \leq \frac{\alpha k_1 + (1-\alpha) k_2}{k_1 \rho_2 (\xi_1)} = \frac{\alpha}{\rho_2 (\xi_1)} + \frac{1-\alpha}{\rho_2 (\xi_2)},$$

and $c_2 (\xi) = 1/\rho_2 (\xi)$ is convex and finite on A. Clearly $\rho_2 (\xi) = 1/c_2 (\xi)$ on A. Setting $c_2 (\xi) = \infty$ for $\xi \in \widetilde{\Delta} \setminus A$, convexity of $c_2 : \widetilde{\Delta} \to (0, \infty]$ is maintained and $\rho_2 (\xi) = 1/c_2 (\xi)$ for all $\xi \in \widetilde{\Delta}$.

Hence, $G_\xi (t)$ has the representation (44) with $\widetilde{\Delta}$ nonempty and convex, $c_1 : \widetilde{\Delta} \to (0, \infty)$ concave, and $c_2 : \widetilde{\Delta} \to (0, \infty]$ convex.

Step 5. By Lemma 42,[29] $\inf_{\xi \in \widetilde{\Delta}} c_1 (\xi) > 0$ and $\inf_{\xi \in \widetilde{\Delta}} c_2 (\xi) > 0$ are necessary and sufficient for the uniform equicontinuity of the family $\{G_\xi (\cdot)\}_{\xi \in \widetilde{\Delta}}$. (Note however that the latter will be a consequence of the fact that $c_2 (\xi)$ is lower semicontinuous over the compact set $\widetilde{\Delta}$.)

Step 6. $\widetilde{\Delta}$ is closed and c_1 is upper semicontinuous on $\widetilde{\Delta}$. The function

$$\xi \mapsto G_\xi (1) = \begin{cases} \frac{1}{c_1(\xi)} & \text{if } \xi \in \widetilde{\Delta} \\ \infty & \text{if } \xi \in \Delta \setminus \widetilde{\Delta} \end{cases}$$

is lower semicontinuous on Δ, and $\inf_{\xi \in \widetilde{\Delta}} c_1 (\xi) > 0$, that is $\lambda = \sup_{\xi \in \widetilde{\Delta}} (1/c_1 (\xi)) < \infty$. Then

$$\widetilde{\Delta} = \{\xi \in \Delta : G_\xi (1) < \infty\} = \{\xi \in \Delta : G_\xi (1) \leq \lambda\}$$

is closed. If $\alpha > 0$,

$$\left\{\xi \in \widetilde{\Delta} : c_1 (\xi) \geq \alpha\right\} = \left\{\xi \in \widetilde{\Delta} : G_\xi (1) \leq \alpha^{-1}\right\}$$

is closed, while if $\alpha \leq 0$ then $\left\{\xi \in \widetilde{\Delta} : c_1 (\xi) \geq \alpha\right\} = \widetilde{\Delta}$. Therefore c_1 is upper semicontinuous.

[29] Remember that in $[0, \infty]$, with $1/0 = \infty$ and $1/\infty = 0$, $a \geq b$ iff $1/a \leq 1/b$, $\inf_{i \in I} a_i = 1/\sup_{i \in I} (1/a_i)$, $\sup_{i \in I} a_i = 1/\inf_{i \in I} (1/a_i)$.

Step 7. $c_2 : \widetilde{\Delta} \to (0, \infty]$ is lower semicontinuous. The function $\xi \mapsto G_\xi(-1) = -1/c_2(\xi)$ is lower semicontinuous (Δ and hence) on $\widetilde{\Delta}$. If $\alpha > 0$,

$$\left\{\xi \in \widetilde{\Delta} : c_2(\xi) \le \alpha\right\} = \left\{\xi \in \widetilde{\Delta} : \frac{1}{c_2(\xi)} \ge \frac{1}{\alpha}\right\} = \left\{\xi \in \widetilde{\Delta} : G_\xi(-1) \le -\frac{1}{\alpha}\right\}$$

is closed, while if $\alpha \le 0$ then $\left\{\xi \in \widetilde{\Delta} : c_2(\xi) \le \alpha\right\} = \emptyset$. Therefore c_2 is lower semicontinuous.

Conversely, Lemma 44 shows that g is finite, monotone, upper semicontinuous, positively homogeneous, quasiconcave. Clearly $\operatorname{dom}(G_\xi) \in \{\emptyset, \mathbb{R}\}$ for all $\xi \in \Delta$. Moreover, $\inf_{\xi\in\widetilde{\Delta}} c_1(\xi) > 0$ (by assumption) and $\inf_{\xi\in\widetilde{\Delta}} c_2(\xi) > 0$ (since c_2 is lower semicontinuous and strictly positive over the compact set $\widetilde{\Delta}$); therefore the family $\{G_\xi(\cdot)\}_{\xi\in\widetilde{\Delta}} = \{G_\xi\}_{\xi\in\Delta:\operatorname{dom}(G_\xi)=\mathbb{R}}$ is uniformly equicontinuous (by Lemma 42). Theorem 38 delivers uniform continuity of g.

Finally, g is normalized if and only if, for all $t \in \mathbb{R}$,

$$t = g(te) = \inf_{\xi\in\Delta} G_\xi(t) = \begin{cases} \inf_{\xi\in\widetilde{\Delta}} \frac{t}{c_1(\xi)} & \text{if } t \ge 0 \\ \inf_{\xi\in\widetilde{\Delta}} \frac{t}{c_2(\xi)} & \text{if } t \le 0 \end{cases} = \begin{cases} t \inf_{\xi\in\widetilde{\Delta}} \frac{1}{c_1(\xi)} & \text{if } t \ge 0 \\ t \sup_{\xi\in\widetilde{\Delta}} \frac{1}{c_2(\xi)} & \text{if } t \le 0 \end{cases}$$

which is equivalent to $\max_{\xi\in\widetilde{\Delta}} c_1(\xi) = \min_{\xi\in\widetilde{\Delta}} c_2(\xi) = 1$ thanks to the semicontinuity properties of c_1 and c_2. ■

A.3 Continuity of Monotone Functionals

A.3.1 Lower and Upper Continuity

Lemma 45 *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 $\{\alpha \in [0,1] : g(\alpha x + (1-\alpha) y) \le c\}$ is closed;

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

Proof. (i) implies (ii). Let $c \in \mathbb{R}$, $S(g,c) = \{x \in Y : g(x) \le c\}$. 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 $x - \varepsilon_m e \le x_{n_m}$ and monotonicity implies $g(x - \varepsilon_m e) \le g(x_{n_m}) \le c$, and left continuity guarantees $g(x) = \lim_m g(x - \varepsilon_m e) \le c$. (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 $g(\alpha_n x + (1-\alpha_n) y) \le c$. 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. $g(\alpha_0 x + (1-\alpha_0) y) \le c$). (This implication does not require lower openness.)

(iii) implies (iv). Let $c \in \mathbb{R}$ and $x, y \in Y$ (with $y \le x$) and $g(x) > c$. Assume, per contra, $g(\alpha x + (1-\alpha) y) \le c$ for all $\alpha \in (0,1)$. By (iii) the set $A = \{\alpha \in [0,1] : g(\alpha x + (1-\alpha) y) \le c\}$ 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 for all $n \geq \bar{n}$

$$x_n \geq x_0 - \alpha_y \varepsilon e = (1-\alpha_y)x_0 + \alpha_y y$$

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 46 *Let X be hyper-Archimedean. For a monotone function $g : X(T) \to \mathbb{R}$, conditions (i)-(iv) of Lemma 45 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 47 *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 46. If T is lower open, then $X(T)$ is lower open too, and Lemma 45 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,

$$T \ni t \leq \operatorname{ess\,inf}((x-\varepsilon_m e) \vee te) \leq \operatorname{ess\,sup}(((x-\varepsilon_m e) \vee te)) = \operatorname{ess\,sup}((x-\varepsilon_m e)) \vee t \leq \operatorname{ess\,sup}(x) \in T,$$

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 $g(x) = \lim_m g(y_m) \leq c$.[30]

(ii) implies (iii) and (iii) implies (iv) are proved in exactly the same way as in Lemma 45.

(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 for all $n \geq \bar{n}$

$$x_n \geq x_0 - \alpha_y \varepsilon e = (1-\alpha_y)x_0 + \alpha_y y$$

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 47, 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 for all $n \geq \bar{n}$

$$x_n \geq x_0 - \alpha_k \varepsilon e = x_0 - \alpha_k \operatorname{ess\,inf}(x_0)e + \alpha_k ke \geq x_0 - \alpha_k x_0 + \alpha_k ke = (1-\alpha_k)x_0 + \alpha_k ke$$

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 47, 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. ■

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.

[30] 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 $c = \lim_n g(x_n) \geq g(x)$.

A.3.2 Uniform Continuity and Lipschitzianity

Proposition 48 *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

$$g(x + \delta e) \leq g(x) + \varepsilon \tag{52}$$

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

$$x - \delta e \leq y \leq x + \delta e. \tag{53}$$

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

$$\begin{aligned}
&t \leq \operatorname{ess\,inf}((x - \delta e) \vee te) \leq \operatorname{ess\,sup}((x - \delta e) \vee te) \leq \operatorname{ess\,sup}(x) \leq \tau, \\
&t \leq \operatorname{ess\,inf}(x) \leq \operatorname{ess\,inf}((x + \delta e) \wedge \tau e) \leq \operatorname{ess\,sup}((x + \delta e) \wedge \tau e) \leq \tau, \\
&t \leq \operatorname{ess\,inf}(x) \leq \operatorname{ess\,inf}(x \vee (t + \delta) e) \leq \operatorname{ess\,sup}(x \vee (t + \delta) e) = \operatorname{ess\,sup}(x) \vee (t + \delta) \leq \tau, \\
&t \leq \operatorname{ess\,inf}(x) \wedge (\tau - \delta) = \operatorname{ess\,inf}(x \wedge (\tau - \delta) e) \leq \operatorname{ess\,sup}(x \wedge (\tau - \delta) e) \leq \operatorname{ess\,sup}(x) \leq \tau,
\end{aligned}$$

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

$$(x \vee (t + \delta) e) - \delta e = (x - \delta e) \vee te \in X(T) \text{ and } (x \wedge (\tau - \delta) e) + \delta e = (x + \delta e) \wedge \tau e \in X(T). \tag{54}$$

From (53) we have $(x - \delta e) \vee te \leq y \leq (x + \delta e) \wedge \tau e$. By monotonicity, (54), and the choice of δ,

$$\begin{aligned} g(x) - \varepsilon &\leq g((x \vee (t+\delta) e)) - \varepsilon \leq g((x \vee (t+\delta) e) - \delta e) = g((x - \delta e) \vee te) \\ &\leq g(y) \leq g((x + \delta e) \wedge \tau e) = g((x \wedge (\tau - \delta) e) + \delta e) \leq g((x \wedge (\tau - \delta) e)) + \varepsilon \\ &\leq g(x) + \varepsilon \end{aligned}$$

and so $g(x) - \varepsilon \leq g(y) \leq g(x) + \varepsilon$, as desired. ■

A similar argument, can be used to prove the following Lipschitz version of Proposition 48

Proposition 49 *A monotone $g : X(T) \longrightarrow \mathbb{R}$ is ℓ-Lipschitz on $X(T)$ if and only if $g(x + \delta e) \leq g(x) + \ell\delta$ for all $x \in X(T)$ and all $\delta > 0$ such that $x + \delta e \in X(T)$.*

Proof.$^{\#}$ If g is ℓ-Lipschitz, then $g(x + \delta e) - g(x) \leq \ell \|x + \delta e - x\| = \ell\delta$ for all $x \in X$ and all $\delta > 0$ such that $x + \delta e \in X(T)$.

Conversely, let $x, y \in X(T)$ with $x \neq y$, and set $t = \operatorname{ess\,inf}(x \wedge y) = \operatorname{ess\,inf}(x) \wedge \operatorname{ess\,inf}(y) \in T$ and $\tau = \operatorname{ess\,sup}(x \vee y) = \operatorname{ess\,sup}(x) \vee \operatorname{ess\,sup}(y) \in T$. From $te \leq x, y \leq \tau e$ it follows $x - y \leq \tau e - te$ and $y - x \leq \tau e - te$. Hence, $|x - y| \leq (\tau - t) e$ and $\|x - y\| \leq \tau - t$. Set $\delta = \|x - y\| > 0$ and notice that $t \leq t + \delta \leq \tau$ and $t \leq \tau - \delta \leq \tau$. Any $z \in X$ such that $te \leq z \leq \tau e$ belongs to $X(T)$. In particular, since

$$\begin{aligned} & t \leq \operatorname{ess\,inf}((x - \delta e) \vee te) \leq \operatorname{ess\,sup}((x - \delta e) \vee te) \leq \operatorname{ess\,sup}(x) \leq \tau, \\ & t \leq \operatorname{ess\,inf}(x) \leq \operatorname{ess\,inf}((x + \delta e) \wedge \tau e) \leq \operatorname{ess\,sup}((x + \delta e) \wedge \tau e) \leq \tau, \\ & t \leq \operatorname{ess\,inf}(x) \leq \operatorname{ess\,inf}(x \vee (t + \delta) e) \leq \operatorname{ess\,sup}(x \vee (t + \delta) e) = \operatorname{ess\,sup}(x) \vee (t + \delta) \leq \tau, \\ & t \leq \operatorname{ess\,inf}(x) \wedge (\tau - \delta) = \operatorname{ess\,inf}(x \wedge (\tau - \delta) e) \leq \operatorname{ess\,sup}(x \wedge (\tau - \delta) e) \leq \operatorname{ess\,sup}(x) \leq \tau, \end{aligned}$$

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

$$(x \vee (t + \delta) e) - \delta e = (x - \delta e) \vee te \in X(T) \text{ and } (x \wedge (\tau - \delta) e) + \delta e = (x + \delta e) \wedge \tau e \in X(T). \quad (55)$$

Since $\|x - y\| \leq \delta$, then $x - \delta e \leq y \leq x + \delta e$, and

$$(x - \delta e) \vee te \leq y \leq (x + \delta e) \wedge \tau e$$

by monotonicity, (55), and the observation that $g(z + \delta' e) \leq g(z) + \ell\delta'$ for all $z \in X(T)$ and all $\delta' > 0$ such that $z + \delta' e \in X(T)$ also implies $g(z - \delta' e) \geq g(z) - \ell\delta'$ for all $z \in X(T)$ and all $\delta' > 0$ such that $z - \delta' e \in X(T)$, it follows that:

$$\begin{aligned} g(x) - \ell\delta &\leq g((x \vee (t+\delta) e)) - \ell\delta \leq g((x \vee (t+\delta) e) - \delta e) = g((x - \delta e) \vee te) \\ &\leq g(y) \leq g((x + \delta e) \wedge \tau e) = g((x \wedge (\tau - \delta) e) + \delta e) \leq g((x \wedge (\tau - \delta) e)) + \ell\delta \\ &\leq g(x) + \ell\delta \end{aligned}$$

and $g(x) - \ell \|x - y\| \leq g(y) \leq g(x) + \ell \|x - y\|$.

Thus $|g(x) - g(y)| \leq \ell \|x - y\|$ for all $x, y \in X(T)$ with $x \neq y$. As wanted. ■

A.3.3 Linear Continuity

Lemma 50 *If $G \in \mathcal{G}(T \times \Delta)$, then*

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

is finite, (evenly) quasiconcave, monotone, normalized, and $G(t, \xi) \geq G_{\xi}(t)$ for all $(t, \xi) \in T \times \Delta$.

Moreover, if $X(T) = B_0(\Sigma, T)$, then g is continuous if and only if G satisfies the following conditions for every partition $A_1, ..., A_n$ of S in Σ, $t_0, t_1, ..., t_n \in T$, and $c \in \mathbb{R}$:

1. *for all* $\varepsilon > 0$ *such that* $G\left(\sum_{i=1}^{n} t_i p(A_i), p\right) > c + \varepsilon$ *for all* $p \in \Delta$*, there exist* $\delta > 0$ *and* $\alpha \in (0,1)$ *such that* $G\left(\sum_{i=1}^{n} (\alpha t_i + (1-\alpha) t_0) p(A_i), p\right) > c + \delta$ *for all* $p \in \Delta$*;*

2. *if there exists* $p \in \Delta$ *such that* $G\left(\sum_{i=1}^{n} t_i p(A_i), p\right) < c$*, then there exist* $q \in \Delta$ *and* $\alpha \in (0,1)$ *such that* $G\left(\sum_{i=1}^{n} (\alpha t_i + (1-\alpha) t_0) q(A_i), q\right) < c$.

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 $x \geq y$, then $\langle \xi, x\rangle \geq \langle \xi, y\rangle$ for all $\xi \in \Delta$, and monotonicity of $G(\cdot, \xi)$ implies that $G(\langle \xi, x\rangle, \xi) \geq G(\langle \xi, y\rangle, \xi)$, and hence $g(x) \geq g(y)$.

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

Finiteness follows from monotonicity and normalization, in fact, for all $x \in X(T)$,

$$\operatorname{ess\,inf}(x)\, e \leq x \leq \operatorname{ess\,sup}(x)\, e \implies \operatorname{ess\,inf}(x) \leq g(x) \leq \operatorname{ess\,sup}(x).$$

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

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

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

Let $\bar{x} \notin \{y \in X(T) : g(y) \geq a\}$, then,

- either $\bar{x} \notin X(T)$ and hence $\bar{x} \notin L$;

- 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 $G\left(\left\langle \bar{\xi}, \bar{x}\right\rangle, \bar{\xi}\right) < a$, and (by monotonicity of G in the first component) for all $y \in X(T)$ such that $\bar{\xi}(y) \leq \bar{\xi}(\bar{x})$

  $$g(y) \leq G\left(\left\langle \bar{\xi}, y\right\rangle, \bar{\xi}\right) \leq G\left(\left\langle \bar{\xi}, \bar{x}\right\rangle, \bar{\xi}\right) < a$$

  that is $\left\{y \in X(T) : \bar{\xi}(y) \leq \bar{\xi}(\bar{x})\right\} \subseteq \{y \in X(T) : g(y) < a\}$ and

  $$\{y \in X(T) : g(y) \geq a\} \subseteq \left\{y \in X(T) : \bar{\xi}(y) > \bar{\xi}(\bar{x})\right\} \subseteq \left[\bar{\xi} > \bar{\xi}(\bar{x})\right];$$

  thus $\left(\bar{\xi}, \bar{\xi}(\bar{x})\right) \in \Delta \times \mathbb{R} : \left[\bar{\xi} > \bar{\xi}(\bar{x})\right] \supseteq \{y \in X(T) : g(y) \geq a\}$ but $\bar{x} \notin \left[\bar{\xi} > \bar{\xi}(\bar{x})\right]$, and hence $\bar{x} \notin L$.

Therefore L is contained $\{y \in X(T) : g(y) \geq a\}$, and the two sets coincide.

Moreover, for all $\left(\bar{t}, \bar{\xi}\right) \in T \times \Delta$, and all $y \in X(T)$ such that $\left\langle \bar{\xi}, y\right\rangle \leq \bar{t}$, then

$$g(y) = \inf_{\xi\in\Delta} G(\langle \xi, y\rangle, \xi) \leq G\left(\left\langle \bar{\xi}, y\right\rangle, \bar{\xi}\right) \leq G\left(\bar{t}, \bar{\xi}\right).$$

Therefore,

$$G_{\bar{\xi}}\left(\bar{t}\right) = \sup_{y\in X(T):\left\langle \bar{\xi}, y\right\rangle\leq\bar{t}} g(y) \leq G\left(\bar{t}, \bar{\xi}\right).$$

Finally, by point (vi) of Proposition 46, g is lower semicontinuous on $X(T) \iff$ for each $t_0 \in T$, $c \in \mathbb{R}$ and $\sum_{i=1}^{n} t_i 1_{A_i} \in X(T)$ with $\inf_{p\in\Delta} G\left(\left\langle p, \sum_{i=1}^{n} t_i 1_{A_i}\right\rangle, p\right) > c$, there exists a number $\alpha \in (0,1)$ such that $\inf_{p\in\Delta} G\left(\left\langle p, \alpha\sum_{i=1}^{n} t_i 1_{A_i} + (1-\alpha) t_0\right\rangle, p\right) > c \iff$ for every partition $A_1, ..., A_n$ of S in Σ, $t_0, t_1, ..., t_n \in T$, and $c \in \mathbb{R}$, $\inf_{p\in\Delta} G\left(\left\langle p, \sum_{i=1}^{n} t_i 1_{A_i}\right\rangle, p\right) > c$ implies that there is $\alpha \in (0,1)$ such that $\inf_{p\in\Delta} G\left(\left\langle p, \alpha\sum_{i=1}^{n} t_i 1_{A_i} + (1-\alpha) t_0\right\rangle, p\right) > c \iff$ for every partition $A_1, ..., A_n$ of S in Σ, $t_0, t_1, ..., t_n \in T$, and $c \in \mathbb{R}$, if there exists $\varepsilon > 0$ such that $G\left(\sum_{i=1}^{n} t_i p(A_i), p\right) > c+\varepsilon$ for all $p \in \Delta$, then there exist $\alpha \in (0,1)$ and $\delta > 0$ such that $G\left(\sum_{i=1}^{n} (\alpha t_i + (1-\alpha) t_0) p(A_i), p\right) > c+\delta$ for all $p \in \Delta$. While, again by Proposition 46, g is upper semicontinuous on $X(T) \iff$ for each $t_0 \in T$, $c \in \mathbb{R}$ and $\sum_{i=1}^{n} t_i 1_{A_i} \in X(T)$ with $g\left(\sum_{i=1}^{n} t_i 1_{A_i}\right) < c$, there is $\alpha \in (0,1)$ such that $g\left(\alpha\sum_{i=1}^{n} t_i 1_{A_i} + (1-\alpha) t_0\right) < c$ $\iff$ for each $t_0 \in T$, $c \in \mathbb{R}$ and $\sum_{i=1}^{n} t_i 1_{A_i} \in X(T)$ with $\inf_{p\in\Delta} G\left(\left\langle p, \sum_{i=1}^{n} t_i 1_{A_i}\right\rangle, p\right) < c$, there is $\alpha \in (0,1)$ such that $\inf_{p\in\Delta} G\left(\left\langle p, \alpha\sum_{i=1}^{n} t_i 1_{A_i} + (1-\alpha) t_0\right\rangle, p\right) < c \iff$ for every partition $A_1, ..., A_n$ of S in Σ, $t_0, t_1, ..., t_n \in T$, and $c \in \mathbb{R}$, $\inf_{p\in\Delta} G\left(\left\langle p, \sum_{i=1}^{n} t_i 1_{A_i}\right\rangle, p\right) < c$ implies that there exists a number $\alpha \in (0,1)$ such that $\inf_{p\in\Delta} G\left(\left\langle p, \alpha\sum_{i=1}^{n} t_i 1_{A_i} + (1-\alpha) t_0\right\rangle, p\right) < c \iff$ for every partition $A_1, ..., A_n$ of S in Σ, $t_0, t_1, ..., t_n \in T$, and $c \in \mathbb{R}$, if there exists $p \in \Delta$ such that $G\left(\sum_{i=1}^{n} t_i p(A_i), p\right) < c$, then there exist $\alpha \in (0,1)$ and $q \in \Delta$ such that $G\left(\sum_{i=1}^{n} (\alpha t_i + (1-\alpha) t_0) q(A_i), q\right) < c$. ■

Lemma 51 *If $G \in \mathcal{H}(T \times \Delta)$, then*

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

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,

$$\begin{aligned} |\langle \xi_\alpha, x_\alpha\rangle - \langle \xi, x\rangle| &\le |\langle \xi_\alpha, x_\alpha\rangle - \langle \xi_\alpha, x\rangle| + |\langle \xi_\alpha, x\rangle - \langle \xi, x\rangle| = |\langle \xi_\alpha, x_\alpha - x\rangle| + |\langle \xi_\alpha, x\rangle - \langle \xi, x\rangle| \\ &\le \|\xi_\alpha\| \, \|x_\alpha - x\| + |\langle \xi_\alpha, x\rangle - \langle \xi, x\rangle| = \|x_\alpha - x\| + |\langle \xi_\alpha, x\rangle - \langle \xi, x\rangle| \to 0. \end{aligned}$$

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

$$g(x) = \inf_{\xi \in \Delta} \Gamma(x, \xi) = \min_{\xi \in \Delta} \Gamma(x, \xi) = \min_{\xi \in \Delta} G(\langle \xi, x \rangle, \xi) \tag{56}$$

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 (56), 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,

$$\liminf_{n} g(x_n) = \lim_{k} g(x_{n_k}) = \lim_{\alpha} g\left(x_{n_{k_\alpha}}\right) = \lim_{\alpha} \Gamma\left(x_{n_{k_\alpha}}, \xi_{n_{k_\alpha}}\right) \geq \Gamma\left(x, \bar{\xi}\right) \geq \min_{\xi \in \Delta} \Gamma(x, \xi) = g(x)$$

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 $\xi \in \Delta$,[31] finally $g(\cdot) = \inf_{\xi \in \Delta} \Gamma(\cdot, \xi)$ is upper semicontinuous too. ■

Lemma 52 *If $G \in \mathcal{E}(T \times \Delta)$, then*

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

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 $\operatorname{dom}(G(\cdot, \xi)) = T$, 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 50), 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 $\operatorname{dom}(G(\cdot, \xi_x)) = T$. 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. ■

[31] 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 $\limsup_n G(\langle \xi, x_n \rangle, \xi) \leq G(\langle \xi, x \rangle, \xi)$.

A.3.4 Monotone Continuity on Function Spaces

Theorem 53 *Let Σ be a σ-algebra, and $I : B_0(\Sigma) \to \mathbb{R}$ be such that*

$$I(\varphi) = \inf_{p\in\Delta} G\left(\int \varphi dp, p\right), \tag{57}$$

where $G : \mathbb{R}\times\Delta \to (-\infty, \infty]$ is jointly lower semicontinuous, grounded,[32] *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 $I\left(k1_{E_n} + \psi 1_{E_n^c}\right) > I(\varphi)$;*
- *(iii) $G(\cdot, p) \equiv \infty$ for all $p \notin \Delta^\sigma$;*
- *(iv) there is $q \in \Delta^\sigma$ such that $\{p \in \Delta : G(t,p) \le \alpha\}$ 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 $p \notin \Delta^\sigma(q)$;*

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) $\displaystyle\bigcup_{t,\alpha\in\mathbb{R}} \{p \in \Delta : G(t,p) \le \alpha\} \subseteq P$;
- (c) $\displaystyle\bigcup_{m,n=1}^{\infty} \{p \in \Delta : G(m,p) \le n\} \subseteq P$.

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

(i) implies (ii). Suppose first that $k \le \min\psi$. 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 $k1_{E_n} + \psi 1_{E_n^c} \ge (\min\psi) 1_{E_n} + \psi 1_{E_n^c}$, 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 $I\left(k1_{E_{n_0}} + \psi 1_{E_{n_0}^c}\right) \ge I\left((\min\psi) 1_{E_{n_0}} + \psi 1_{E_{n_0}^c}\right) > I(\varphi)$.

(ii) implies (iii). By the Claim, it is enough to show that $\{p \in \Delta : G(t,p) \le \alpha\} \subseteq \Delta^\sigma$ for all t and α in $\mathbb{R}$. Let $E_n \downarrow \emptyset$ and $r \in \{p \in \Delta : G(t,p) \le \alpha\}$. Set $\varphi \equiv \alpha$ and $\psi \equiv \beta$ with $\beta > \alpha \vee 0$. For each $k > 0$ there is $n_k \ge 1$ such that $\alpha = I(\varphi) < I\left(-k1_{E_{n_k}} + \beta 1_{E_{n_k}^c}\right)$.[33] Thus, since

$$-k1_{E_{n_k}} + \beta 1_{E_{n_k}^c} \le -k1_{E_n} + \beta 1_{E_n^c}, \qquad \forall n \ge n_k.$$

[32] That is, such that $\inf_{p\in\Delta} G(t,p) = t$ for all $t \in \mathbb{R}$.
[33] The first equality descends from the normalization of I that corresponds to the groundedness of G.

it follows

$$\alpha < I\left(-k1_{E_n} + \beta 1_{E_n^c}\right) = \inf_{p\in\Delta} G\left(\left\langle p, -k1_{E_n} + \beta 1_{E_n^c}\right\rangle, p\right), \qquad \forall n \geq n_k.$$

If $\left\langle r, -k1_{E_n} + \beta 1_{E_n^c}\right\rangle \leq t$ for some $n \geq n_k$, monotonicity of $G(\cdot, r)$ would deliver

$$\inf_{p\in\Delta} G\left(\left\langle p, -k1_{E_n} + \beta 1_{E_n^c}\right\rangle, p\right) \leq G\left(\left\langle r, -k1_{E_n} + \beta 1_{E_n^c}\right\rangle, r\right) \leq G(t, r) \leq \alpha$$

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

$$\begin{aligned} -kr(E_n) + \beta(1 - r(E_n)) &> t, \qquad \forall n \geq n_k \\ -kr(E_n) + \beta &> t, \qquad \forall n \geq n_k \\ r(E_n) &< \frac{\beta - t}{k}, \qquad \forall n \geq n_k \end{aligned}$$

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., [20, Prop. 2.13]). Then, for all $m, n \in \mathbb{N}$, there is $q_{(n,m)} \in \Delta^\sigma$ such 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 [11, Rem. 5.5], Γ-converges) to γ_0. By [11, 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

$$g(x) = \int_\Delta \phi(\langle \xi, x\rangle)\, d\mu(\xi), \qquad \forall x \in X, \tag{58}$$

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 (58) we need some notation. We denote by $ca(\mathcal{B}(\Delta))$ 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. Finally,

$$ba(\mathcal{B}(\Delta), \mu) = \{\nu \in ba(\mathcal{B}(\Delta)) : B \in \mathcal{B}(\Delta) \text{ and } \mu(B) = 0 \text{ implies } \nu(B) = 0\}$$

is (isometrically isomorphic to, e.g., [49, 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 $\alpha \in \mathbb{R}$,

$$\begin{aligned} A(\alpha x + y)(\xi) &= \langle \xi, \alpha x + y \rangle = \alpha \langle \xi, x \rangle + \langle \xi, y \rangle \\ &= \alpha (Ax)(\xi) + Ay(\xi) = (\alpha Ax + Ay)(\xi) \qquad \forall \xi \in \Delta, \end{aligned}$$

hence $A(\alpha x + y) = \alpha Ax + Ay$. A is bounded, in fact,

$$|Ax(\xi)| = |\langle \xi, x \rangle| \le \|\xi\| \|x\| = \|x\| \qquad \forall \xi \in \Delta, x \in X,$$

thus

$$\|Ax\|_{L^{\infty}(\mu)} \le 1 \|x\| \qquad \forall x \in X,$$

and $|||A||| \le 1$. 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

$$\langle A^{*}\nu, x \rangle = \langle \nu, Ax \rangle = \int_{\Delta} Ax d\nu = \int_{\Delta} \langle \xi, x \rangle d\nu(\xi), \qquad \forall x \in X. \tag{59}$$

A^{*} is continuous and $A^{*}\nu$ is denoted by $\int_{\Delta} \xi d\nu(\xi)$ in view of (59).

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 every $\xi \in X_{+}^{*}$, define

$$\begin{aligned} \Gamma(\xi) &= (A^{*})^{-1}(\xi) \cap ca_{+}(\mathcal{B}(\Delta), \mu) = \{\nu \in ca_{+}(\mathcal{B}(\Delta), \mu) : A^{*}\nu = \xi\} \\ &= \left\{\nu \in ca_{+}(\mathcal{B}(\Delta), \mu) : \int_{\Delta} \zeta d\nu(\zeta) = \xi\right\}. \end{aligned}$$

$\Gamma(\xi)$ is a (possibly empty) closed and convex (hence weakly closed) subset of $ca_{+}(\mathcal{B}(\Delta), \mu)$ and $\nu(\Delta) = \xi(e)$ for all $\nu \in \Gamma(\xi)$.[34]

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,

[34] $(A^{*})^{-1}(\xi)$ 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 $\nu \in \Gamma(\xi)$, $\nu \ge 0$ and $A^{*}\nu = \xi$, hence $\nu(\Delta) = \langle A^{*}\nu, e \rangle = \xi(e)$.

- 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 54 *The functional (58) is finite, concave, continuous and monotone on X.*

Its conjugate is, for all $\xi \in X^*$,

$$g^*(\xi) = \sup\left\{\int_\Delta \phi^*\left(\frac{d\nu}{d\mu}(\zeta)\right)d\mu(\zeta) : \nu \in \Gamma(\xi)\right\}.$$

with the convention $g^*(\xi) = -\infty$ if $\Gamma(\xi) = \emptyset$.

Moreover, for all $(t,\xi) \in \mathbb{R}\times\Delta$,

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

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 [42] and [43].

By [43, 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

$$I_\phi^*(\nu) = I_{\phi^*}(u^*) = \int_\Delta \phi^*(u^*(\zeta))\,d\mu(\zeta) \tag{60}$$

if there exists $u^* \in L^1(\mu)$ such that

$$\nu(u) = \int_\Delta u(\zeta)\,u^*(\zeta)\,d\mu(\zeta), \qquad \forall u \in L^\infty(\mu),$$

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).[35] Therefore,

$$I_\phi^*(\nu) = \begin{cases} \int_\Delta \phi^*\left(\frac{d\nu}{d\mu}(\zeta)\right)d\mu(\zeta) & \text{if } \nu \text{ is countably additive,} \\ -\infty & \text{otherwise.} \end{cases} \tag{61}$$

Consider the bounded linear operator

$$\begin{array}{rccc} A: & X & \to & L^\infty(\mu) \\ & x & \mapsto & \langle\cdot,x\rangle \end{array}$$

[35] 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 $u^* = d\nu/d\mu$.

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 $L^\infty(\mu)$, [43, 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

$$A^* I_\phi^*(\xi) = \sup\left\{I_\phi^*(\nu) : \nu \in ba(\mathcal{B}(\Delta), \mu),\ A^*\nu = \xi\right\}. \tag{62}$$

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 (62) implies

$$g^*(\xi) = \sup\left\{I_\phi^*(\nu) : \nu \in ba_+(\mathcal{B}(\Delta), \mu),\ A^*\nu = \xi\right\}. \tag{63}$$

By (61), $I_\phi^*(\nu) = -\infty$ for all $\nu \notin ca(\mathcal{B}(\Delta), \mu)$. Then (63) amounts to

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

and (61) 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 Lemma 31, $G_\xi = g_\xi$ for all $\xi \in \Delta$. By Corollary 36, for each $(t, \xi) \in \mathbb{R} \times \Delta$,

$$G_\xi(t) = \inf_{k \geq 0}\{kt - g^*(k\xi)\} = \inf_{k \geq 0}\left\{tk - \sup\left\{\int_\Delta \phi^*\left(\frac{d\nu}{d\mu}(\zeta)\right) d\mu(\zeta) : \nu \in \Gamma(k\xi)\right\}\right\},$$

thus, if $\Gamma(\xi) \neq \emptyset$, it follows that

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

else, if $\Gamma(\xi) = \emptyset$,

$$\begin{aligned}
&\sup\left\{\int_\Delta \phi^*\left(\tfrac{d\nu}{d\mu}(\zeta)\right) d\mu(\zeta) : \nu \in \Gamma(k\xi)\right\} = \sup\left\{\int_\Delta \phi^*\left(\tfrac{d(k\gamma)}{d\mu}(\zeta)\right) d\mu(\zeta) : \gamma \in \Gamma(\xi)\right\} = -\infty \quad \text{if } k > 0 \\
&\sup\left\{\int_\Delta \phi^*\left(\tfrac{d\nu}{d\mu}(\zeta)\right) d\mu(\zeta) : \nu \in \Gamma(k\xi)\right\} = \phi^*(0) = \inf_{k \in \mathbb{R}}\{-\phi(k)\} = -\sup_{k \in \mathbb{R}} \phi(k) \quad \text{if } k = 0
\end{aligned}$$

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

$$g(x) = \phi^{-1}\left[\int_\Delta \phi(\langle \xi, x\rangle) d\mu(\xi)\right] \tag{64}$$

of (58). 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

$$\widehat{\phi^{-1}}(t) = \begin{cases} \infty & \text{if } t \geq a \\ \phi^{-1}(t) & \text{if } a > t > -\infty \\ -\infty & \text{if } t = -\infty \end{cases}, \tag{65}$$

this extension is simply denoted ϕ^{-1}. Application of Theorem 54 and Lemma 32 delivers, for all $(t,\xi) \in \mathbb{R}\times\Delta$,

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

$$G_\xi(t) = \phi^{-1}\left(\inf\left\{\inf_{k\geq 0}\left[tk - \int_\Delta \phi^*\left(k\frac{d\nu}{d\mu}(\zeta)\right) d\mu(\zeta)\right] : \nu \in \Gamma(\xi)\right\}\right) \tag{66}$$

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

Lemma 55 *For a twice differentiable $\phi : \mathbb{R} \to \mathbb{R}$ with $\phi' > 0$ and $\phi'' < 0$, the following facts are equivalent:*

- *(i) $J_\lambda(\cdot) = \phi\left(\phi^{-1}(\cdot) + \lambda\right)$ 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.

Proof. ϕ^{-1} is differentiable too with strictly positive derivative. Setting $\phi^{-1}(r) = \psi(r)$, we get

$$\begin{aligned} J_\lambda(r) &= \phi[\psi(r) + \lambda] \\ \psi'(r) &= \frac{1}{\phi'(\psi(r))} \end{aligned}$$

and J_λ is twice differentiable with

$$\begin{aligned} J'_\lambda(r) &= \frac{\phi'[\psi(r)+\lambda]}{\phi'(\psi(r))} \\ J''_\lambda(r) &= \frac{\dfrac{\phi''[\psi(r)+\lambda]}{\phi'(\psi(r))}\phi'(\psi(r)) - \phi'[\psi(r)+\lambda]\dfrac{\phi''(\psi(r))}{\phi'(\psi(r))}}{[\phi'(\psi(r))]^2} \end{aligned}$$

for all $r \in \phi(\mathbb{R}), \lambda \geq 0$. Therefore (i) is equivalent to

$$\begin{aligned} &\frac{\dfrac{\phi''(\psi(r)+\lambda)}{\phi'(\psi(r))}\phi'(\psi(r)) - \phi'(\psi(r)+\lambda)\dfrac{\phi''(\psi(r))}{\phi'(\psi(r))}}{[\phi'(\psi(r))]^2} \leq 0 \quad \forall r \in \phi(\mathbb{R}), \lambda \geq 0 \Leftrightarrow \\ &\phi''(\psi(r)+\lambda)\phi'(\psi(r)) - \phi'(\psi(r)+\lambda)\phi''(\psi(r)) \leq 0 \quad \forall r \in \phi(\mathbb{R}), \lambda \geq 0 \Leftrightarrow \\ &\frac{\phi''(\psi(r)+\lambda)\phi'(\psi(r))}{\phi''(\psi(r))} - \phi'(\psi(r)+\lambda) \geq 0 \quad \forall r \in \phi(\mathbb{R}), \lambda \geq 0 \Leftrightarrow \\ &\frac{\phi'(\psi(r))}{\phi''(\psi(r))} - \frac{\phi'(\psi(r)+\lambda)}{\phi''(\psi(r)+\lambda)} \leq 0 \quad \forall r \in \phi(\mathbb{R}), \lambda \geq 0 \Leftrightarrow \\ &-\frac{\phi'(\psi(r)+\lambda)}{\phi''(\psi(r)+\lambda)} \leq -\frac{\phi'(\psi(r))}{\phi''(\psi(r))} \quad \forall r \in \phi(\mathbb{R}), \lambda \geq 0 \end{aligned}$$

which amounts to (ii) since ψ is onto. ■

Proposition 56 *If the scalar functions $J_\lambda(r) = \phi\left[\phi^{-1}(r) + \lambda\right]$ are concave on $\phi(\mathbb{R})$ for all $\lambda \geq 0$, then (64) is 1-Lipschitz.*

Proof. By the Jensen Inequality, we have

$$\begin{aligned} J_\lambda\left(\int_\Delta \phi(\langle \xi, x\rangle)\, d\mu(\xi)\right) &\geq \int_\Delta J_\lambda(\phi(\langle \xi, x\rangle))\, d\mu(\xi) \\ \phi\left(\phi^{-1}\left(\int_\Delta \phi(\langle \xi, x\rangle)\, d\mu(\xi)\right) + \lambda\right) &\geq \int_\Delta \phi(\langle \xi, x\rangle + \lambda)\, d\mu(\xi) \\ \phi^{-1}\left(\int_\Delta \phi(\langle \xi, x\rangle)\, d\mu(\xi)\right) + \lambda &\geq \phi^{-1}\left(\int_\Delta \phi(\langle \xi, x + \lambda e\rangle)\, d\mu(\xi)\right) \\ g(x) + \lambda &\geq g(x + \lambda e) \end{aligned} \tag{67}$$

for all $\lambda \geq 0$. Proposition 49 delivers 1-Lipschitzianity. ■

Proposition 57 *The functional (64) 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

$$\phi^{-1}\left(\frac{\phi(t+\lambda) + \phi(r+\lambda)}{2}\right) = \phi^{-1}\left(\frac{\phi(t) + \phi(r)}{2}\right) + \lambda \qquad \forall t, r, \lambda \in \mathbb{R}. \tag{68}$$

By [14, p. 28] ϕ is CARA.

We report his argument for the sake of completeness. Wlog, assume $\phi(0) = 0$. Next observe that J_λ is affine for all $\lambda \in \mathbb{R}$, in fact it is continuous and, for all $t' = \phi(t), r' = \phi(r) \in \phi(\mathbb{R})$, by (68)

$$\begin{aligned} J_\lambda\left(\frac{t'+r'}{2}\right) &= \phi\left[\phi^{-1}\left(\frac{t'+r'}{2}\right) + \lambda\right] = \phi\left[\phi^{-1}\left(\frac{\phi(t)+\phi(r)}{2}\right) + \lambda\right] \\ &= \phi\left[\phi^{-1}\left(\frac{\phi(t+\lambda)+\phi(r+\lambda)}{2}\right)\right] = \frac{\phi(t+\lambda)+\phi(r+\lambda)}{2} \\ &= \frac{\phi\left(\phi^{-1}(\phi(t)) + \lambda\right)}{2} + \frac{\phi\left(\phi^{-1}(\phi(r)) + \lambda\right)}{2} \\ &= \frac{\phi\left(\phi^{-1}(t') + \lambda\right)}{2} + \frac{\phi\left(\phi^{-1}(r') + \lambda\right)}{2} = \frac{1}{2}J_\lambda(t') + \frac{1}{2}J_\lambda(r'). \end{aligned}$$

Moreover, $J_\lambda(0) = \phi\left(\phi^{-1}(0) + \lambda\right) = \phi(\lambda)$ for all $\lambda \in \mathbb{R}$. It follows that there exists $k : \mathbb{R} \to \mathbb{R}$ such that

$$J_\lambda(w) = k(\lambda)\, w + \phi(\lambda) \qquad \forall w \in \phi(\mathbb{R}), \lambda \in \mathbb{R}.$$

Again by (68), for all $t, r, \lambda \in \mathbb{R}$,

$$\frac{\phi(t+\lambda)+\phi(r+\lambda)}{2} = \phi\left(\phi^{-1}\left(\frac{\phi(t)+\phi(r)}{2}\right)+\lambda\right) = J_\lambda\left(\frac{\phi(t)+\phi(r)}{2}\right) = k(\lambda)\frac{\phi(t)+\phi(r)}{2}+\phi(\lambda) \tag{69}$$

thus, for $r=0$, $\phi(t+\lambda)+\phi(\lambda) = k(\lambda)\phi(t)+2\phi(\lambda)$, or

$$\phi(t+\lambda) = k(\lambda)\phi(t)+\phi(\lambda)$$

and exchanging the roles of t and λ

$$\phi(\lambda+t) = k(t)\phi(\lambda)+\phi(t) \tag{70}$$

hence

$$k(\lambda)\phi(t)+\phi(\lambda) = k(t)\phi(\lambda)+\phi(t) \tag{71}$$

thus if $t, \lambda \neq 0$

$$\frac{k(\lambda)-1}{\phi(\lambda)} = \frac{k(t)-1}{\phi(t)} = C$$

that is there exists a constant C such that

$$k(t) = C\phi(t)+1 \tag{72}$$

for all $t \neq 0$, but also if $t=0$, (71) delivers $k(0)=1$, and (72) holds.

Finally, plugging (72) in (70),

$$\phi(\lambda+t) = \phi(\lambda)+\phi(t)+C\phi(\lambda)\phi(t) \qquad \forall t, \lambda \in \mathbb{R}.$$

If $C=0$, ϕ is linear; else

$$\begin{aligned}(C\phi(t)+1)(C\phi(\lambda)+1) &= C\phi(t)C\phi(\lambda)+C\phi(\lambda)+C\phi(t)+1\\ &= C(\phi(\lambda)+\phi(t)+C\phi(\lambda)\phi(t))+1\\ &= C\phi(t+\lambda)+1.\end{aligned}$$

Thus $C\phi(t)+1$ is exponential. ■

B.1.1 Relative Entropy

Here we further study the CARA case.

Proposition 58 *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 $(t,\xi) \in \mathbb{R}\times\Delta$,

$$\begin{aligned}g^*(\xi) &= -\frac{1}{\theta}\inf\{R(\nu \parallel \mu) : \nu \in \Gamma(\xi)\}, \\ G_\xi(t) &= t+\frac{1}{\theta}\inf\{R(\nu \parallel \mu) : \nu \in \Gamma(\xi)\}.\end{aligned} \tag{73}$$

Proof. We first consider the case $\theta = 1$. In view of Theorem 54, 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 and increasing. Next we evaluate $\phi^*(t)$. Set $\psi(t) = e^t$,

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

Since $\phi(t) = -\psi(-t)$, then

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

Claim. For all $\nu \in \Delta^\sigma(\mu)$ and $t \in \mathbb{R}$

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

Proof of the Claim. First,

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

If $R(\nu \parallel \mu) < \infty$, then, for all $k \geq 0$,

$$\begin{aligned}
\int_\Delta \left(k\frac{d\nu}{d\mu}(\zeta) - k\frac{d\nu}{d\mu}(\zeta)\log k\frac{d\nu}{d\mu}(\zeta)\right)d\mu(\zeta) &= \int_\Delta \left(k\frac{d\nu}{d\mu}(\zeta) - k\frac{d\nu}{d\mu}(\zeta)\log k - k\frac{d\nu}{d\mu}(\zeta)\log\frac{d\nu}{d\mu}(\zeta)\right)d\mu(\zeta) \\
&= \int_\Delta (k - k\log k)\frac{d\nu}{d\mu}(\zeta) - k\frac{d\nu}{d\mu}(\zeta)\log\frac{d\nu}{d\mu}(\zeta)\,d\mu(\zeta) \\
&= k - k\log k - kR(\nu \parallel \mu) \\
&= k - k\log k - k\log e^{R(\nu\|\mu)} \\
&= k - k\log ke^{R(\nu\|\mu)} \\
&= \frac{1}{e^{R(\nu\|\mu)}}\left(e^{R(\nu\|\mu)}k - e^{R(\nu\|\mu)}k\log e^{R(\nu\|\mu)}k\right) \\
&= \frac{1}{e^{R(\nu\|\mu)}}\phi^*\left(e^{R(\nu\|\mu)}k\right),
\end{aligned}$$

and

$$\begin{aligned}
\inf_{k\geq 0}\left[tk - \int_\Delta \phi^*\left(k\frac{d\nu}{d\mu}(\zeta)\right)d\mu(\zeta)\right] &= \inf_{k\geq 0}\left[tk - \frac{1}{e^{R(\gamma\|\mu)}}\phi^*\left(e^{R(\gamma\|\mu)}k\right)\right] \\
&= \frac{1}{e^{R(\gamma\|\mu)}}\inf_{k\geq 0}\left[t\left(e^{R(\gamma\|\mu)}k\right) - \phi^*\left(e^{R(\gamma\|\mu)}k\right)\right] \\
&= \frac{1}{e^{R(\gamma\|\mu)}}\inf_{r\geq 0}\left[tr - \phi^*(r)\right] = \frac{1}{e^{R(\gamma\|\mu)}}\inf_{r\in\mathbb{R}}\left[tr - \phi^*(r)\right] \\
&= \frac{1}{e^{R(\gamma\|\mu)}}\phi^{**}(t) = \frac{1}{e^{R(\gamma\|\mu)}}\phi(t) = -e^{-t}e^{-R(\nu\|\mu)}.
\end{aligned}$$

Else if $R(\nu \parallel \mu) = \infty$

$$\int_\Delta \left(k\frac{d\nu}{d\mu}(\zeta) - k\frac{d\nu}{d\mu}(\zeta)\log k\frac{d\nu}{d\mu}(\zeta)\right)d\mu(\zeta) = \int_\Delta (k - k\log k)\frac{d\nu}{d\mu}(\zeta) - k\frac{d\nu}{d\mu}(\zeta)\log\frac{d\nu}{d\mu}(\zeta)\,d\mu(\zeta)$$

is 0 if $k = 0$ and $-\infty$ otherwise. Then

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

As wanted. $\square$

As a consequence, in view of Theorem 54, 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,

$$\begin{aligned} G_{\xi}(t) &= -\log\left(e^{-t}\sup\left\{e^{-R(\nu\|\mu)} : \nu \in \Gamma(\xi)\right\}\right) \\ &= -\log e^{-t} - \log\left(\sup\left\{e^{-R(\nu\|\mu)} : \nu \in \Gamma(\xi)\right\}\right) \\ &= t - \sup\left\{\log e^{-R(\nu\|\mu)} : \nu \in \Gamma(\xi)\right\} \\ &= t - \sup\left\{-R(\nu \parallel \mu) : \nu \in \Gamma(\xi)\right\} \\ &= t + \inf\left\{R(\nu \parallel \mu) : \nu \in \Gamma(\xi)\right\} \end{aligned}$$

while, if $\Gamma(\xi)$ is empty, then

$$G_{\xi}(t) = -\log(0) = \infty = t + \inf\left\{R(\nu \parallel \mu) : \nu \in \Gamma(\xi)\right\}.$$

Monotonicity, translation invariance, concavity, and finiteness of g are easily shown. The conjugate of g can then be calculated by (iii) of Corollary 37, thus for all $\xi \in \Delta$, $g^{*}(\xi) = -g_{\xi}(0) = -\inf\left\{R(\nu \parallel \mu) : \nu \in \Gamma(\xi)\right\}$.

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 37. $\blacksquare$

C A Family of Statistical Distance Functions

Throughout this section we adopt the convention $0\cdot\infty = 0/0 = 0$. We consider a strictly increasing concave function $\phi : \mathbb{R} \to \mathbb{R}$ (with $\sup_{\mathbb{R}}\phi = a$) and a countably additive probability measure μ on a measurable space endowed with a σ-algebra $\mathcal{B}$ that contains at least two singletons (e.g. the σ-algebra $\mathcal{B}(\Delta)$ considered in the previous section, provided Δ contains at least two elements). We extend ϕ by continuity to $[-\infty,\infty]$, by setting $\phi(-\infty) = -\infty$ and $\phi(\infty) = a$ and we extend ϕ^{-1} (again by continuity) as in (65). It is important to notice that such functions are extended valued continuous and monotone.

We extend $I_t(\cdot \parallel \mu)$, as defined by (19), to $\Delta(\mathcal{B},\mu)$ by setting $I_t(\gamma \parallel \mu) = \infty$ if $\gamma \notin \Delta^{\sigma}(\mathcal{B},\mu)$. Before proving the basic properties of $I_t(\cdot \parallel \mu)$, it is worth noticing few facts. The function

$$g(f) = \int \phi(f)\,d\mu, \qquad \forall f \in L^{\infty}(\mu),$$

is (finite) concave, continuous, and monotone, see ([43, Corollary 2A]). It is well known that $(L^\infty(\mu), \|\cdot\|_\infty, \geq)$ is a normed Riesz space with order unit, $\|\cdot\|_\infty$ is its supnorm, and its topological dual is $ba(\mathcal{B},\mu)$. In particular, see again ([43, Corollary 2A]), for all $\gamma \in \Delta(\mathcal{B},\mu)$ and all $k \geq 0$,

$$g^*(k\gamma) = \begin{cases} \int \phi^*\left(k\frac{d\gamma}{d\mu}\right) d\mu & \text{if } k = 0 \text{ or } \gamma \in \Delta^\sigma(\mathcal{B},\mu), \\ -\infty & \text{otherwise.} \end{cases}$$

By Lemma 31 and Corollary 36, for all $(t,\gamma) \in \mathbb{R} \times \Delta(\mathcal{B},\mu)$,

$$G_\gamma(t) = g_\gamma(t) = \inf_{k\in\mathbb{R}_+} \{kt - g^*(k\gamma)\} = \begin{cases} \inf_{k\geq 0}\left\{kt - \int \phi^*\left(k\frac{d\gamma}{d\mu}\right) d\mu\right\} & \text{if } \gamma \in \Delta^\sigma(\mathcal{B},\mu), \\ \sup_{\mathbb{R}} \phi & \text{otherwise.} \end{cases}$$

By Lemmas 28 and 29 the mapping $(t,\gamma) \mapsto G_\gamma(t)$ is quasiconvex and lower semicontinuous (when $\mathbb{R}$ is endowed with the usual topology and $\Delta(\mathcal{B},\mu)$ is endowed with the weak* topology). From

$$I_t(\gamma \parallel \mu) = \phi^{-1}(G_\gamma(t)) - t, \qquad \forall (t,\gamma) \in \mathbb{R} \times \Delta(\mathcal{B},\mu), \tag{74}$$

we obtain some important properties of $I_t(\cdot \parallel \mu)$.

Proof of Proposition 11. Indeed we show that for all $t \in \mathbb{R}$,

(i) $I_t(\mu \parallel \mu) = 0$;

(ii) $I_t(\gamma \parallel \mu) \geq 0$ for each $\gamma \in \Delta(\mathcal{B},\mu)$;

(iii) $I_t(\cdot \parallel \mu)$ is quasiconvex, weak* lower semicontinuous on $\Delta(\mathcal{B},\mu)$, and $\{\gamma \in \Delta(\mathcal{B},\mu) : I_t(\gamma \parallel \mu) \leq c\}$ is a weakly compact subset of $\Delta^\sigma(\mathcal{B},\mu)$ for all $c \in \mathbb{R}$.

Monotonicity of ϕ guarantees that $\operatorname{dom} \phi^* \subseteq \mathbb{R}_+$.

(i) By the Fenchel-Moreau Theorem,

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

(ii) The inequality is trivial if $\gamma \notin \Delta^\sigma(\mathcal{B},\mu)$. Else, by the Jensen inequality, for all $k \geq 0$,

$$\int \phi^*\left(k\frac{d\gamma}{d\mu}\right) d\mu \leq \phi^*\left(\int k\frac{d\gamma}{d\mu} d\mu\right) = \phi^*(k) = \int \phi^*\left(k\frac{d\mu}{d\mu}\right) d\mu.$$

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

(iii) Let $c \in \mathbb{R}$ and set,

$$C = \{\nu \in \Delta(\mathcal{B},\mu) : I_t(\nu \parallel \mu) \leq c\}.$$

Next we show $C = \{\nu \in \Delta(\mathcal{B},\mu) : G_\nu(t) \leq \phi(c+t)\}$.

$\subseteq$) Let $\gamma \in C$, then positivity of $I_t(\cdot \parallel \mu)$ guarantees that

$$t \leq \phi^{-1}(G_\gamma(t)) \leq c + t,$$

and $G_\gamma(t) \in \phi(\mathbb{R})$, therefore

$$\phi(t) \leq G_\gamma(t) \leq \phi(c+t).$$

$\supseteq$) Let γ be such that $G_\gamma(t) \leq \phi(c+t)$, monotonicity of ϕ^{-1} delivers $\phi^{-1}(G_\gamma(t)) \leq c + t$, that is $I_t(\gamma \parallel \mu) \leq c$.

By Lemmas 28 and 29 the mapping $(t,\gamma)\mapsto G_{\gamma}(t)$ is quasiconvex and lower-semicontinuous. Therefore $\{\nu\in\Delta(\mathcal{B},\mu):G_{\nu}(t)\leq\phi(c+t)\}$, that is C, is convex and weak* compact. The observation that C consists of countably additive measures delivers weak compactness (e.g., [20, Prop. 2.13]). ■

Proof of Proposition 13. We first prove that (i) implies (ii). Let $\mu\in\Delta^{\sigma}(\mathcal{B})$ and $t\in\mathbb{R}$. The inequality holds by definition if $I_t^2(\gamma\parallel\mu)=\infty$. Assume $I_t^2(\gamma\parallel\mu)=c<\infty$, then

$$\phi_2(t)\leq G_{\gamma}^2(t)\leq\phi_2(c+t). \tag{75}$$

Since ϕ_1 is more concave than ϕ_2, then, for each $f\in L^{\infty}(\mu)$, by Jensen's inequality,

$$\int\phi_1(f)\,d\mu\leq h\left(\int\phi_2(f)\,d\mu\right),$$

hence,

$$G_{\gamma}^1(t)=\sup_{\int fd\gamma\leq t}\int\phi_1(f)\,d\mu\leq\sup_{\int fd\gamma\leq t}h\left(\int\phi_2(f)\,d\mu\right)\leq h\left(G_{\gamma}^2(t)\right).$$

Where the last inequality descends from (75) and monotonicity of h. Moreover, notice that $G_{\gamma}^2(t)\in\phi_2(\mathbb{R})$, therefore $h\left(G_{\gamma}^2(t)\right)\in\phi_1(\mathbb{R})$, and so does $G_{\gamma}^1(t)$. Finally,

$$I_t^1(\gamma\parallel\mu)+t=(\phi_1)^{-1}\left(G_{\gamma}^1(t)\right)\leq(\phi_1)^{-1}\left(h\left(G_{\gamma}^2(t)\right)\right)=(\phi_2)^{-1}\left(G_{\gamma}^2(t)\right)=I_t^2(\gamma\parallel\mu)+t.$$

Conversely, let $\mu\in\Delta^{\sigma}(\mathcal{B})$. The function $h=\phi_1\circ(\phi_2)^{-1}:\phi_2(\mathbb{R})\to\mathbb{R}$ is strictly increasing and $\phi_1=h\circ\phi_2$. Then, for any $f\in L^{\infty}(\mu)$,

$$h\left(\int\phi_2(f)\,d\mu\right)=h\left(\min_{\gamma\in\Delta(\mathcal{B},\mu)}G_{\gamma}^2\left(\int fd\gamma\right)\right)=\left(\phi_1\circ(\phi_2)^{-1}\right)\left(G_{\bar{\gamma}}^2\left(\int fd\bar{\gamma}\right)\right), \tag{76}$$

where $\bar{\gamma}\in\Delta^{\sigma}(\mathcal{B},\mu)$ and $G_{\bar{\gamma}}^2\left(\int fd\bar{\gamma}\right)\in\phi_2(\mathbb{R})$. Set $\bar{t}=\int fd\bar{\gamma}$.

This implies that $\phi_2^{-1}\left(G_{\bar{\gamma}}^2(\bar{t})\right)\in\mathbb{R}$ and by (74)

$$\phi_2^{-1}\left(G_{\bar{\gamma}}(\bar{t})\right)=I_{\bar{t}}^2(\bar{\gamma}\parallel\mu)+\bar{t}.$$

This and (ii) yield $0\leq I_{\bar{t}}^1(\bar{\gamma}\parallel\mu)\leq I_{\bar{t}}^2(\bar{\gamma}\parallel\mu)<\infty$. By (76), we conclude that

$$h\left(\int\phi_2(f)\,d\mu\right)=\phi_1\left(I_{\bar{t}}^2(\bar{\gamma}\parallel\mu)+\bar{t}\right)\geq\phi_1\left(I_{\bar{t}}^1(\bar{\gamma}\parallel\mu)+\bar{t}\right).$$

Since $I_{\bar{t}}^1(\bar{\gamma}\parallel\mu)=\phi_1^{-1}\left(G_{\bar{\gamma}}^1(\bar{t})\right)-\bar{t}$ is finite, then $G_{\bar{\gamma}}^1(\bar{t})\in\phi_1(\mathbb{R})$ and

$$\begin{aligned}\phi_1\left(I_{\bar{t}}^1(\bar{\gamma}\parallel\mu)+\bar{t}\right)&=\phi_1\left(\phi_1^{-1}\left(G_{\bar{\gamma}}^1(\bar{t})\right)-\bar{t}+\bar{t}\right)=G_{\bar{\gamma}}^1(\bar{t})\\&=G_{\bar{\gamma}}^1\left(\int fd\bar{\gamma}\right)\geq\min_{\gamma\in\Delta(\mathcal{B},\mu)}G_{\gamma}^1\left(\int fd\gamma\right)=\int\phi_1(f)\,d\mu=\int h(\phi_2(f))\,d\mu.\end{aligned}$$

Finally, $h\left(\int\phi_2(f)\,d\mu\right)\geq\int h(\phi_2(f))\,d\mu$ for all $\mu\in\Delta^{\sigma}(\mathcal{B})$ and all $f\in L^{\infty}(\mu)$. Since $\mathcal{B}$ contains two singletons, this implies that h is concave. ■

Corollary 59 *Let $\phi_1,\phi_2:\mathbb{R}\to\mathbb{R}$ two strictly increasing and concave functions, then the following conditions are equivalent:*

(i) ϕ_1 is a positive affine transformation of ϕ_2 (i.e. $\phi_1\approx\phi_2$);

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

C.1 Order Orlicz Functions

Lemma 60 *If ϕ is order Orlicz, then $\lim_{t\to\infty}\phi'(t)=0$ and $\lim_{t\to-\infty}\phi'(t)=\infty$.*

Next proposition regroups the properties of functions that satisfy these tail conditions.

Proposition 61 *Let $\phi:\mathbb{R}\to\mathbb{R}$ be strictly increasing, strictly concave, differentiable, with $\phi(0)=0$, $\phi'(0)=1$, $\lim_{t\to\infty}\phi'(t)=0$, $\lim_{t\to-\infty}\phi'(t)=\infty$, and set $\psi=(\phi')^{-1}$.*

- $\lim_{t\to\infty}\phi(t)/t=0$ *and* $\lim_{t\to-\infty}\phi(t)/t=\infty$.
- $\psi:(0,\infty)\to(-\infty,\infty)$ *is continuous and strictly decreasing with* $\lim_{t\to 0^+}\psi(t)=\infty$ *and* $\lim_{t\to\infty}\psi(t)=-\infty$.
- ϕ^* *is strictly concave on* $(0,\infty)$ *and*

$$\phi^*(t)=\begin{cases}\min_{k\in\mathbb{R}}(kt-\phi(k))=t\psi(t)-\phi(\psi(t)) & \text{if } t>0,\\ -\sup_{k\in\mathbb{R}}\phi(k) & \text{if } t=0,\\ -\infty & \text{if } t<0;\end{cases}$$

  moreover, it is differentiable on $(0,\infty)$ *and* $(\phi^*)'=\psi$.

Finally, $\arg\max\phi^*=\{1\}$, $\max\phi^*=0$ *and* ϕ^* *is strictly increasing on* $(0,1)$ *and strictly decreasing on* $(1,\infty)$.

The proofs are long but standard exercises in Convex Analysis (see, e.g., [41]), that we leave to the reader. We extend ψ from $[0,\infty]$ to $[-\infty,\infty]$ by continuity, that is we set $\psi(0)=\infty$ and $\psi(\infty)=-\infty$. This delivers

$$\phi^*(0)=-\sup_{k\in\mathbb{R}}\phi(k)=-\phi(\infty)=-\phi(\psi(0))=0\psi(0)-\phi(\psi(0)).$$

That is $\phi^*(t)=t\psi(t)-\phi(\psi(t))$ for all $t\geq 0$ and $-\infty$ otherwise.

Next proposition shows the effects that the constraints on the elasticity of ϕ in the definition of order Orlicz impose on its conjugate ϕ^*. This is a variation on classical results in the theory of Orlicz spaces (see, e.g., [31]).

Proposition 62 *Let $\phi:\mathbb{R}\to\mathbb{R}$ be strictly increasing, strictly concave, differentiable, with $\phi(0)=0$, $\phi'(0)=1$, $\lim_{t\to\infty}\phi'(t)=0$, $\lim_{t\to-\infty}\phi'(t)=\infty$, and set $\rho=-\phi^*$.*

The following statements are equivalent:

i) *There exists $T_1>1>\varepsilon_1>0$ and $h\in\mathbb{R}_{++}$ such that $\rho(k/2)\leq h\rho(k)$ for each $k\in(0,\varepsilon_1)$ and $\rho(2k)\leq h\rho(k)$ for each $k\in(T_1,\infty)$.*

ii) *There exists $T_2>1>\varepsilon_2>0$ such that for each $l\in(1,\infty)$ there exists $h(l)\in\mathbb{R}_{++}$ such that for each $k\in(T_2,\infty)$*

$$\rho(lk)\leq h(l)\rho(k)$$

and for each $l\in(0,1)$ there exists $h(l)\in\mathbb{R}_{++}$ such that for each $k\in(0,\varepsilon_2)$

$$\rho(lk)\leq h(l)\rho(k).$$

iii) *There exists $T_3>1>\varepsilon_3>0$ and $\alpha\in(1,\infty)$ such that for each $k\in(0,\varepsilon_3)\cup(T_3,\infty)$,*

$$\left|\frac{k\rho'(k)}{\rho(k)}\right|\leq\alpha. \tag{77}$$

iv) There exists $T_4 > 0 > t_4$ and $\alpha \in (1, \infty)$ such that for each $t \in (T_4, \infty)$

$$\frac{t\phi'(t)}{\phi(t)} \leq \frac{\alpha}{(\alpha+1)},$$

and for each $t \in (-\infty, t_4)$

$$\frac{t\phi'(t)}{\phi(t)} \geq \frac{\alpha}{(\alpha-1)}.$$

That is, ϕ is order Orlicz.

Proof. i)⇒ii). If $l \in (1, \infty)$ then there exists $n \in \mathbb{N}$ such that $2^n \geq l$. Pick $k > T_1$. It follows that $2^n k \geq lk \geq k > T_1$ and $\rho(lk) \leq \rho(2^n k)$. Therefore, since $k \in (T_1, \infty)$, $2^m k \in (T_1, \infty)$ for each $m \in \mathbb{N}$ and we can apply the inequality of i), obtaining $\rho(lk) \leq \rho(2^n k) \leq h\rho\left(2^{n-1}k\right) \leq h^n \rho(k)$. Similarly, if $l \in (0,1)$ then there exists $n \in \mathbb{N}$ such that $2^{-n} \leq l$. Pick $k < \varepsilon_1$. It follows that $2^{-n}k \leq lk \leq k < \varepsilon_1$ and $\rho(lk) \leq \rho(2^{-n}k)$. Therefore, since $k \in (0, \varepsilon_1)$, $2^{-m}k \in (0, \varepsilon_1)$ for each $m \in \mathbb{N}$ and we can apply the inequality of i), obtaining $\rho(lk) \leq \rho(2^{-n}k) \leq h\rho\left(2^{-n+1}k\right) \leq h^n \rho(k)$. If we define $\varepsilon_2 = \varepsilon_1$ and $T_2 = T_1$ the statement is proved.

ii)⇒iii). Pick $l = 2$ then for $k \in (T_2, \infty)$, we have $\rho(2k) \leq h(2)\rho(k)$. Since $T_2 > 1$, $\frac{\rho(2k)}{\rho(k)} > 1$ hence $h(2) > 1$. This implies that for $k \in (T_2, \infty)$,

$$k\rho'(k) = \int_k^{2k} \rho'(k)\, ds \leq \int_k^{2k} \rho'(s)\, ds = \rho(2k) - \rho(k) \leq \rho(2k) \leq h(2)\rho(k).$$

We can conclude that for $k \in (T_3, \infty)$, where $T_3 = T_2$, $|k\rho'(k)/\rho(k)| \leq h(2)$.

Now, pick $l = 1/2$, then for $k \in (0, \varepsilon_2)$, we have $\rho(k/2) \leq h(1/2)\rho(k)$ that in turn implies that for $k \in (0, \varepsilon_2)$,

$$\frac{k}{2}\rho'(k) = \int_{\frac{k}{2}}^{k} \rho'(k)\, ds \geq \int_{\frac{k}{2}}^{k} \rho'(s)\, ds = \rho(k) - \rho\left(\frac{k}{2}\right) \geq -\rho\left(\frac{k}{2}\right) \geq -h\left(\frac{1}{2}\right)\rho(k).$$

This implies that for each $k \in (0, \varepsilon_3)$, where $\varepsilon_3 = \varepsilon_2$, $-k\rho'(k)/\rho(k) = |k\rho'(k)/\rho(k)| \leq 2h(1/2)$.

If we define $\alpha = \max\{2h(1/2), h(2)\}$ then $\alpha \in (1, \infty)$ and we finally obtain that

$$\left|\frac{k\rho'(k)}{\rho(k)}\right| \leq \alpha \qquad \forall k \in (0, \varepsilon_3) \cup (T_3, \infty).$$

iii)⇒iv). By Proposition 61, recall that, for each $k \in (0, \infty)$, $\rho(k) = -k\psi(k) + \phi(\psi(k))$ and $\rho'(k) = -\psi(k)$, where $\psi = (\phi')^{-1}$. By (77), it follows that for each $k \in (T_3, \infty)$, $k\psi(k)(\alpha - 1) \leq \alpha\phi(\psi(k))$.

Set $t_4 = \psi(T_3)$, since $\psi((0, \infty)) = \mathbb{R}$, ψ is strictly decreasing and continuous, and $\psi(1) = (\phi^*)'(1) = 0$, then $t_4 < 0$ and for each $t \in (-\infty, t_4)$ there exists $k \in (T_3, \infty)$ such that $t = \psi(k) = (\phi')^{-1}(k)$, therefore

$$t\phi'(t)(\alpha - 1) = \psi(k)\, k(\alpha - 1) \leq \alpha\phi(\psi(k)) = \alpha\phi(t).$$

Since $t < 0$ and $\alpha > 1$, this implies $t\phi'(t)/\phi(t) \geq \alpha/(\alpha - 1)$ for each $t \in (-\infty, t_4)$.

Similarly, by (77), for $k \in (0, \varepsilon_3)$, $k\psi(k) \leq \alpha[-k\psi(k) + \phi(\psi(k))]$, which implies that for $k \in (0, \varepsilon_3)$, $k\psi(k)(\alpha + 1) \leq \alpha\phi(\psi(k))$.

Set $T_4 = \psi(\varepsilon_3)$, then $T_4 > 0$ and for each $t \in (T_4, \infty)$ there exists $k \in (0, \varepsilon_3)$ such that $t = \psi(k) = (\phi')^{-1}(k)$, therefore

$$t\phi'(t)(\alpha + 1) = \psi(k)\, k(\alpha + 1) \leq \alpha\phi(\psi(k)) = \alpha\phi(t).$$

Since $t>0$ and $\alpha>1$, this implies $t\phi'(t)/\phi(t)\leq\alpha/(\alpha+1)$ for each $t\in(T_4,\infty)$.

iv)⇒i) Let $\varepsilon_1=\phi'(T_4)$. Since $\phi':(-\infty,\infty)\to(0,\infty)$ is onto, strictly decreasing, $\phi'(0)=1$ and $\psi=(\phi')^{-1}$, then $\varepsilon_1\in(0,1)$ and for all $k\in(0,\varepsilon_1)$ there exists $t\in(T_4,\infty)$ such that $k=\phi'(t)$. Therefore $t=\psi(k)$ and $t\phi'(t)(\alpha+1)\leq\alpha\phi(t)$ implies $k\psi(k)(\alpha+1)\leq\alpha\phi(\psi(k))$, that is $-k\rho'(k)/\rho(k)\leq\alpha$. Similarly, $T_1=\phi'(t_4)$ belongs to $(1,\infty)$ and $k\rho'(k)/\rho(k)\leq\alpha$ for all $k\in(T_1,\infty)$.

Thus, for each $k\in(T_1,\infty)$,

$$\log\frac{\rho(2k)}{\rho(k)}=\int_k^{2k}\frac{\rho'(s)}{\rho(s)}ds\leq\alpha\int_k^{2k}\frac{1}{s}ds=\alpha\log 2,$$

which implies for each $k\in(T_1,\infty)$, $\rho(2k)\leq 2^{\alpha}\rho(k)$. Similarly, if $k\in(0,\varepsilon_1)$,

$$\log\frac{\rho(k)}{\rho\left(\frac{k}{2}\right)}=\int_{\frac{k}{2}}^{k}\frac{\rho'(s)}{\rho(s)}ds\geq-\alpha\int_{\frac{k}{2}}^{k}\frac{1}{s}ds=-\alpha\log 2.$$

This implies that $\rho(k)\geq 2^{-\alpha}\rho(k/2)$ for each $k\in(0,\varepsilon_1)$, hence the statement. ■

Remark 63 *Notice that if i) holds, for each $l\in(0,\infty)$ there exists $h_1(l)>0$ such that $\rho(lk)\leq h_1(l)\rho(k)$ for each $k\in\left(T_1\vee l^{-1},\infty\right)$. Indeed, pick $l\in(0,\infty)$, then there exists an $\bar{n}\in\mathbb{N}$ such that $2^{\bar{n}}\geq l$. Since $k\in\left(T_1\vee l^{-1},\infty\right)$ then $2^n k\in\left(T_1\vee l^{-1},\infty\right)$ for each $n\in\mathbb{N}$ and*

$$\rho(lk)\leq\rho\left(2^{\bar{n}}k\right)\leq h^{\bar{n}}\rho(k).$$

Similarly, for each $l\in(0,\infty)$ there exists $h_2(l)>0$ such that $\rho(lk)\leq h_2(l)\rho(k)$ for each $k\in\left(0,\varepsilon_1\wedge l^{-1}\right)$. Indeed, pick $l\in(0,\infty)$, then there exists an $\bar{n}\in\mathbb{N}$ such that $1/2^{\bar{n}}\leq l$. Since $k\in\left(0,\varepsilon_1\wedge l^{-1}\right)$ then $k/2^n\in\left(0,\varepsilon_1\wedge l^{-1}\right)$ for each $n\in\mathbb{N}$ and

$$\rho(lk)\leq\rho\left(2^{-\bar{n}}k\right)\leq h^{\bar{n}}\rho(k).$$

Let $\gamma\in\Delta^{\sigma}(\mathcal{B},\mu)$, for all $t\in\mathbb{R}$, define $F_t:[0,\infty)\to[-\infty,\infty]$ by $k\mapsto kt-\int\phi^*\left(k\frac{d\gamma}{d\mu}\right)d\mu$.

Proposition 64 *Let ϕ be order Orlicz, if $\gamma\in\Delta^{\sigma}(\mathcal{B},\mu)$ then* $\operatorname{int}\operatorname{dom}F_t\in\{\emptyset,\mathbb{R}_{++}\}$.

Proof. If $\operatorname{int}\operatorname{dom}F_t=\emptyset$ there is nothing to prove. Else there exists $k\in\mathbb{R}_{++}$ such that $k\in\operatorname{int}\operatorname{dom}F_t$, and so $\int\rho\left(k\frac{d\gamma}{d\mu}\right)d\mu<\infty$. Fix $l\in(0,\infty)$, define $f=k\frac{d\gamma}{d\mu}$ and call $A=\left\{\omega:f(\omega)>T_1\vee\frac{1}{l}\right\}$, $B=\left\{\omega:f(\omega)\in\left(0,\varepsilon_1\wedge\frac{1}{l}\right)\right\}$, $C=\{\omega:f(\omega)=0\}$ and $D=(A\cup B\cup C)^c$. Then, it follows that

$$\begin{aligned}F_t(lk)&=lkt-\int\phi^*(lf)\,d\mu=lkt+\int\rho(lf)\,d\mu=lkt+\int_A\rho(lf)\,d\mu+\int_B\rho(lf)\,d\mu+\int_C\rho(lf)\,d\mu+\int_D\rho(lf)\,d\mu\\&\leq lkt+h_1(l)\int_A\rho(f)\,d\mu+h_2(l)\int_B\rho(f)\,d\mu+\int_C\rho(f)\,d\mu+(\rho(l\varepsilon_1\wedge 1)\vee\rho(lT_1\vee 1))\,\mu(D)<\infty\end{aligned}$$

■

Indeed Proposition 64 shows that for an order Orlicz ϕ and $\gamma\in\Delta^{\sigma}(\mathcal{B},\mu)$, $\phi^*(r(d\gamma/d\mu))\in L^1(\mu)$ for *some* $r>0$ if and only if $\phi^*(r(d\gamma/d\mu))\in L^1(\mu)$ for *all* $r>0$.

Lemma 65 *Let ϕ be order Orlicz. If $\gamma\in\Delta^{\sigma}(\mathcal{B},\mu)$ and $\phi^*(d\gamma/d\mu)\in L^1(\mu)$, then $\psi(k(d\gamma/d\mu))\in L^1(\gamma)$ for each $k>0$.*

Proof. Notice that for each $h_1, h_2, t > 0$ we have that

$$\psi\left(\left(h_1+h_2\right)t\right)t \leq \frac{\phi^*\left(\left(h_1+h_2\right)t\right)-\phi^*\left(h_1 t\right)}{h_2} \leq \psi\left(h_1 t\right)t. \tag{78}$$

Let $k > 0$. Define $A = \left\{\omega : k\frac{d\gamma}{d\mu}(\omega) = 0\right\}$, $B = \left\{\omega : k\frac{d\gamma}{d\mu}(\omega) \in (1,\infty)\right\}$ and $C = \left\{\omega : k\frac{d\gamma}{d\mu}(\omega) \in (0,1]\right\}$. Set $h_1 = k$ and take any $h_2 \in (0,\infty)$. Then for each $\omega \in B$, since $\psi\left(k\frac{d\gamma}{d\mu}(\omega)\right) \leq 0$ and $\frac{d\gamma}{d\mu}(\omega) > 0$, by (78), we have that

$$\frac{\phi^*\left(\left(k+h_2\right)\frac{d\gamma}{d\mu}(\omega)\right)-\phi^*\left(k\frac{d\gamma}{d\mu}(\omega)\right)}{h_2} \leq \psi\left(k\frac{d\gamma}{d\mu}(\omega)\right)\frac{d\gamma}{d\mu}(\omega) \leq 0 \tag{79}$$

that is

$$0 \leq 1_B\left|\psi\left(k\frac{d\gamma}{d\mu}\right)\right|\frac{d\gamma}{d\mu} \leq \frac{\phi^*\left(k\frac{d\gamma}{d\mu}\right)-\phi^*\left(\left(k+h_2\right)\frac{d\gamma}{d\mu}\right)}{h_2}1_B$$

By Proposition 64, $\phi^*\left(r\left(d\gamma/d\mu\right)\right) \in L^1(\mu)$ for each $r > 0$, therefore, $\int_B\left|\psi\left(k\frac{d\gamma}{d\mu}\right)\right|d\gamma < \infty$.

Consider again (78), but set $h_1 = h_2 = \frac{k}{2}$. If $\omega \in C$, it follows that,

$$0 \leq \psi\left(k\frac{d\gamma}{d\mu}(\omega)\right)\frac{d\gamma}{d\mu}(\omega) \leq \frac{\phi^*\left(k\frac{d\gamma}{d\mu}(\omega)\right)-\phi^*\left(\frac{k}{2}\frac{d\gamma}{d\mu}(\omega)\right)}{k/2}$$

that is

$$0 \leq 1_C\left|\psi\left(k\frac{d\gamma}{d\mu}\right)\right|\frac{d\gamma}{d\mu} \leq \frac{\phi^*\left(k\frac{d\gamma}{d\mu}\right)-\phi^*\left(\frac{k}{2}\frac{d\gamma}{d\mu}\right)}{k/2}1_C$$

By Proposition 64, $\phi^*\left(r\left(d\gamma/d\mu\right)\right) \in L^1(\mu)$ for each $r > 0$, therefore, $\int_C\left|\psi\left(k\frac{d\gamma}{d\mu}\right)\right|d\gamma < \infty$.

Finally $\gamma(A) = 0$, therefore, $\int_A\left|\psi\left(k\frac{d\gamma}{d\mu}\right)\right|d\gamma < \infty$. Conclude that

$$\int\left|\psi\left(k\frac{d\gamma}{d\mu}\right)\right|d\gamma = \int_A\left|\psi\left(k\frac{d\gamma}{d\mu}\right)\right|d\gamma + \int_B\left|\psi\left(k\frac{d\gamma}{d\mu}\right)\right|d\gamma + \int_C\left|\psi\left(k\frac{d\gamma}{d\mu}\right)\right|d\gamma < \infty$$

and $\psi\left(k\left(d\gamma/d\mu\right)\right) \in L^1(\gamma)$. ■

Proof of Proposition 15. By Proposition 61, $\psi : (0,\infty) \to (-\infty,\infty)$ is strictly decreasing and onto.

Let $t \in \mathbb{R}$ and $\gamma \in \Delta^\sigma(\mathcal{B},\mu)$ be such that $I_t(\gamma \,\|\, \mu) < \infty$. It follows that $\phi^*\left(k\frac{d\gamma}{d\mu}\right) \in L^1(\mu)$ for some $k > 0$ and that

$$\inf_{k\geq 0}\left[kt - \int\phi^*\left(k\frac{d\gamma}{d\mu}\right)d\mu\right] = \inf_{k>0}\left[kt - \int\phi^*\left(k\frac{d\gamma}{d\mu}\right)d\mu\right].$$

By Proposition 64 and Lemma 65, we can define $\Gamma : (0,\infty) \to \mathbb{R}$ by $\Gamma(k) = \int\psi\left(k\frac{d\gamma}{d\mu}\right)d\gamma$. In order to show that $k(\gamma) \in (0,\infty)$ is well defined and unique, we will prove that Γ is strictly decreasing and onto. Let $k, h \in (0,\infty)$ such that $k > h$, then $k\frac{d\gamma}{d\mu} > h\frac{d\gamma}{d\mu}$ γ-a.s. and, since ψ is strictly decreasing, we have $\psi\left(h\frac{d\gamma}{d\mu}\right) > \psi\left(k\frac{d\gamma}{d\mu}\right)$ γ-a.s.. By Lemma 65, $\psi\left(h\frac{d\gamma}{d\mu}\right), \psi\left(k\frac{d\gamma}{d\mu}\right) \in L^1(\gamma)$ and it follows that $\Gamma(h) > \Gamma(k)$. Hence, Γ is strictly decreasing.

By the Monotone Convergence Theorem, it follows that

$$\Gamma(1) - \Gamma(n) = \int\left[\psi\left(\frac{d\gamma}{d\mu}\right) - \psi\left(n\frac{d\gamma}{d\mu}\right)\right]d\gamma \to \infty \text{ as } n \to \infty,$$

and that

$$\Gamma\left(\frac{1}{n}\right) - \Gamma(1) = \int\left[\psi\left(\frac{1}{n}\frac{d\gamma}{d\mu}\right) - \psi\left(\frac{d\gamma}{d\mu}\right)\right]d\gamma \to \infty \text{ as } n \to \infty.$$

Monotonicity of Γ implies that

$$\lim_{k\to\infty}\Gamma(k)=-\infty \text{ and } \lim_{k\to 0^+}\Gamma(k)=\infty. \tag{80}$$

By the Dominated Convergence Theorem, for each $k_0\in(0,\infty)$,

$$0\le\Gamma\left(k_0-\frac{1}{n}\right)-\Gamma(k_0)=\int\left[\psi\left(\left(k_0-\frac{1}{n}\right)\frac{d\gamma}{d\mu}\right)-\psi\left(k_0\frac{d\gamma}{d\mu}\right)\right]d\gamma\to 0 \text{ as } n\to\infty,$$

and

$$0\le\Gamma(k_0)-\Gamma\left(k_0+\frac{1}{n}\right)=\int\left[\psi\left(k_0\frac{d\gamma}{d\mu}\right)-\psi\left(\left(k_0+\frac{1}{n}\right)\frac{d\gamma}{d\mu}\right)\right]d\gamma\to 0 \text{ as } n\to\infty.$$

Monotonicity of Γ implies that

$$\lim_{k\to k_0^-}\Gamma(k)=\Gamma(k_0)=\lim_{k\to k_0^+}\Gamma(k). \tag{81}$$

From (81), we derive the continuity of Γ that matched with (80) implies that Γ is surjective. Therefore, for each $t\in\mathbb{R}$ there exists $k^*\in(0,\infty)$ such that $\Gamma(k^*)=t$. Since Γ is strictly decreasing, such k^* is unique. This proves that $k(\gamma)$ exists and it is unique.

$F_t(k)=kt-\int\phi^*\left(k\frac{d\gamma}{d\mu}\right)d\mu$ for $k\in(0,\infty)$ is (finite) convex and differentiable. Convexity is obvious. Next we show differentiability. Let $h,k\in(0,\infty)$, then for each $\omega\notin N=\left\{\omega:\frac{d\gamma}{d\mu}(\omega)=0\right\}$,

$$\psi\left((k+h)\frac{d\gamma}{d\mu}(\omega)\right)\frac{d\gamma}{d\mu}(\omega)\le\frac{\phi^*\left((k+h)\frac{d\gamma}{d\mu}(\omega)\right)-\phi^*\left(k\frac{d\gamma}{d\mu}(\omega)\right)}{h}\le\psi\left(k\frac{d\gamma}{d\mu}(\omega)\right)\frac{d\gamma}{d\mu}(\omega), \tag{82}$$

by (78). Moreover,

$$\int_N\frac{\phi^*\left((k+h)\frac{d\gamma}{d\mu}\right)-\phi^*\left(k\frac{d\gamma}{d\mu}\right)}{h}d\mu=0. \tag{83}$$

This is obvious if $\mu(N)=0$ or if $\phi^*(0)$ is finite; else $\int\phi^*\left(k\frac{d\gamma}{d\mu}\right)d\mu=\int_N\phi^*\left(k\frac{d\gamma}{d\mu}\right)d\mu+\int_{N^c}\phi^*\left(k\frac{d\gamma}{d\mu}\right)d\mu$ $=-\infty$, which is absurd. From (82) and (83), we obtain

$$\Gamma(k+h)\le\int\frac{\phi^*\left((k+h)\frac{d\gamma}{d\mu}\right)-\phi^*\left(k\frac{d\gamma}{d\mu}\right)}{h}d\mu\le\Gamma(k),\qquad\forall h,k\in(0,\infty),$$

which, together with continuity of Γ, delivers $F_t'(r)=t-\Gamma(r)$ for all $r\in(0,\infty)$.

Finally, $F_t'(k)=0$ if and only if $\int\psi\left(k\frac{d\gamma}{d\mu}\right)d\gamma=t$, that is $k=k(\gamma)$ and

$$\inf_{k\ge 0}F_t(k)=\inf_{k>0}F_t(k)=F_t(k(\gamma)).$$

By Proposition 61, we have that $\phi^*(k)=k\psi(k)-(\phi\circ\psi)(k)$ for all $k\ge 0$, and so

$$\begin{aligned}
I_t(\gamma\parallel\mu)&=\phi^{-1}(F_t(k(\gamma)))-t=\phi^{-1}\left(k(\gamma)t-\int\phi^*\left(k(\gamma)\frac{d\gamma}{d\mu}\right)d\mu\right)-t\\
&=\phi^{-1}\left(k(\gamma)t-\int\left(k(\gamma)\frac{d\gamma}{d\mu}\psi\left(k(\gamma)\frac{d\gamma}{d\mu}\right)-(\phi\circ\psi)\left(k(\gamma)\frac{d\gamma}{d\mu}\right)\right)d\mu\right)-t\\
&=\phi^{-1}\left(k(\gamma)t-\int k(\gamma)\frac{d\gamma}{d\mu}\psi\left(k(\gamma)\frac{d\gamma}{d\mu}\right)d\mu+\int(\phi\circ\psi)\left(k(\gamma)\frac{d\gamma}{d\mu}\right)d\mu\right)-t\\
&=\phi^{-1}\left(k(\gamma)t-k(\gamma)\int\psi\left(k(\gamma)\frac{d\gamma}{d\mu}\right)d\gamma+\int(\phi\circ\psi)\left(k(\gamma)\frac{d\gamma}{d\mu}\right)d\mu\right)-t\\
&=\phi^{-1}\left(\int(\phi\circ\psi)\left(k(\gamma)\frac{d\gamma}{d\mu}\right)d\mu\right)-t,
\end{aligned}$$

as desired. ■

D Proofs

Proof of Lemma 1. It is a direct consequence of Lemma 51. ■

Lemma 66 *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 $\bar{\alpha} f + (1-\bar{\alpha}) g \prec g$. 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

$$g \precsim \frac{\bar{\alpha}}{\beta}(\beta f + (1-\beta) g) + \left(1 - \frac{\bar{\alpha}}{\beta}\right) g = \bar{\alpha} f + \frac{\bar{\alpha}}{\beta} g - \bar{\alpha} g + g - \frac{\bar{\alpha}}{\beta} g = \bar{\alpha} f + (1-\bar{\alpha}) g \prec g$$

a contradiction. We conclude that $\alpha f + (1-\alpha) g \succsim g$ for all $\alpha \in (0,1)$, as desired. ■

Lemma 67 *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*

$$f \succsim g \Longleftrightarrow I(u(f)) \geq I(u(g)). \tag{84}$$

Moreover, u is cardinally unique, and, given u, there is a unique normalized $I : B_0(\Sigma, u(X)) \to \mathbb{R}$ that satisfies (84).

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 (84) can be derived using the same technique of [34, Lemma 28], where for the existence of u we use axiom A.4 in place of the stronger Weak Certainty Independence axiom of [34]. In particular, $B_0(\Sigma, u(X)) = \{u(f) : f \in \mathcal{F}\}$.

By Lemma 66, $\succsim$ is a convex preference, and so I is quasiconcave. Continuity follows from A.5 and Proposition 46.

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 (84). 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 68 *Let $\succsim$, I, and u be like in Lemma 67. 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 $f, g \in \mathcal{F}$

$$\frac{1}{2}f(s) + \frac{1}{2}y' \sim \frac{1}{2}g(s) + \frac{1}{2}y \quad \forall s \in S \Longrightarrow \frac{1}{2}x_f + \frac{1}{2}z' \precsim \frac{1}{2}x_g + \frac{1}{2}z. \tag{85}$$

(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 (85) 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

$$\frac{1}{2}f(s) + \frac{1}{2}y' = \frac{1}{2}y + \frac{1}{2}y' \sim \frac{1}{2}y' + \frac{1}{2}y = \frac{1}{2}g(s) + \frac{1}{2}y \quad \forall s \in S$$

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 48, 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

$$u(f(s)) \leq u(g(s)) + u(y) - u(y') \qquad \forall s \in S. \tag{86}$$

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 (86) 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 clearly $\|((\psi \wedge (k - \delta')) + \delta') - (\psi \wedge (k - \delta'))\| = \delta' \leq \delta$ 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 69 *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.* Assume $u(X) = \mathbb{R}$, we want to show that there are $x \succ y$ in X such that, for each $\alpha \in (0, 1)$, there exist $z, z' \in X$ such that $\alpha z + (1 - \alpha) y \succ x \succ y \succ \alpha z' + (1 - \alpha) x$. Let $x \in u^{-1}(1)$, $y \in u^{-1}(-1)$, and for all $\alpha \in (0, 1)$ choose $z = z(\alpha) \in u^{-1}\left(\frac{3}{\alpha}\right)$ and $z' = z'(\alpha) \in u^{-1}\left(-\frac{3}{\alpha}\right)$, to obtain

$$\begin{aligned}
u(\alpha z + (1 - \alpha) y) &= \alpha u(z) + (1 - \alpha) u(y) = 3 - 1 + \alpha \geq 2 \\
u(x) &= 1 \\
u(y) &= -1 \\
u(\alpha z' + (1 - \alpha) x) &= \alpha u(z') + (1 - \alpha) u(x) = -3 + (1 - \alpha) \leq -2.
\end{aligned}$$

Conversely, assume there are $x \succ y$ in X such that, for each $\alpha \in (0, 1)$, there exist $z, z' \in X$ such that $\alpha z + (1 - \alpha) y \succ x \succ y \succ \alpha z' + (1 - \alpha) x$. Wlog, assume $u(x) = 1$ and $u(y) = -1$. For all $n \in \mathbb{N}$ there exist $z_n, z'_n \in X$ such that

$$\begin{aligned}
&\frac{1}{n}z_n + \left(1 - \frac{1}{n}\right) y \succ x \succ y \succ \frac{1}{n}z'_n + \left(1 - \frac{1}{n}\right) x, \text{ i.e.} \\
&\frac{1}{n}u(z_n) - 1 + \frac{1}{n} > 1 > -1 > \frac{1}{n}u(z'_n) + 1 - \frac{1}{n}
\end{aligned}$$

then $u(z_n) > 2n - 1$ and $1 - 2n > u(z'_n)$ for all $n \in \mathbb{N}$. Thus $u(X)$ cannot be bounded above or bounded below, and therefore it coincides with $\mathbb{R}$. ■

Proof of Theorem 3. Suppose (i) holds, i.e., $\succsim$ satisfies axioms A.1-A.5. By Lemma 67, 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

$$f \succsim g \Longleftrightarrow I(u(f)) \geq I(u(g)).$$

By Corollary 35, $I(\varphi) = \inf_{p \in \Delta} G_p(\langle p, \varphi \rangle)$ for all $\varphi \in B_0(\Sigma, u(X))$, i.e.,

$$I(u(f)) = \inf_{p \in \Delta} G_p\left(\int u(f)\,dp\right) \qquad \forall f \in \mathcal{F},$$

where $G_p(t) = \sup\{I(\varphi) : \varphi \in B_0(\Sigma, u(X)) \text{ and } \langle p, \varphi \rangle \leq t\}$ for all $(t, p) \in u(X) \times \Delta$.[36]

Lemma 28 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 $t \in u(X)$,

$$t = I(t) = \inf_{p \in \Delta} G_p(\langle p, t \rangle) = \inf_{p \in \Delta} G_p(t).$$

[36] 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 $\varphi \in B_0(\Sigma, u(X))$.

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 (7) holds. This proves (ii).

Conversely, suppose (ii) holds. Since $G \in \mathcal{G}(u(X) \times \Delta)$, then, by Lemma 50,

$$I(\varphi) = \inf_{p \in \Delta} G(\langle p, \varphi \rangle, p) \qquad \forall \varphi \in B_0(\Sigma, u(X)), \tag{87}$$

is finite, (evenly) quasiconcave, monotone, normalized. Linear continuity of G implies continuity of I, and (7) amounts to

$$f \succsim g \Longleftrightarrow I(u(f)) \geq I(u(g)). \tag{88}$$

Lemma 67 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 $\mathcal{F}$,

$$f \succsim g \Longleftrightarrow \inf_{p \in \Delta} H\left(\int v(f)\, dp, p\right) \geq \inf_{p \in \Delta} H\left(\int v(g)\, dp, p\right). \tag{89}$$

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

$$J(\varphi) = \inf_{p \in \Delta} H(\langle p, \varphi \rangle, p) \qquad \forall \varphi \in B_0(\Sigma, v(X)). \tag{90}$$

Since $H \in \mathcal{G}(v(X) \times \Delta)$, then, by Lemma 50, J is finite, (evenly) quasiconcave, monotone, normalized,

$$H(t,p) \geq \sup\{J(\varphi) : \varphi \in B_0(\Sigma, v(X)) \text{ and } \langle p, \varphi \rangle \leq t\} \qquad \forall (t,p) \in v(X) \times \Delta, \tag{91}$$

and (89) amounts to

$$f \succsim g \Longleftrightarrow J(v(f)) \geq J(v(g)). \tag{92}$$

Since J is normalized, by (92), v represents $\succsim$ on X, then it is cardinally equivalent to u. Assume $v = u$, then (88), (92), and Lemma 67 guarantee that $J = I$ (in particular H is linearly continuous too). By (91), 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 35...") 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

$$f \succsim g \Longleftrightarrow \inf_{p \in \Delta} G^\star\left(\int u(f)\, dp, p\right) \geq \inf_{p \in \Delta} G^\star\left(\int u(g)\, dp, p\right). \tag{93}$$

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 $(t,p) \in u(X) \times \Delta$,

$$\begin{aligned}
\sup_{f \in \mathcal{F}} \left\{u(x_f) : \int u(f)\, dp \leq t\right\} &= \sup_{f \in \mathcal{F}} \left\{I(u(f)) : \int u(f)\, dp \leq t\right\} \\
&= \sup\{I(\varphi) : \varphi \in B_0(\Sigma, u(X)) \text{ and } \langle p, \varphi \rangle \leq t\} \\
&= G_p(t) = G^\star(t,p).
\end{aligned}$$

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 29, 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 (93) 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 (8), for all $(t,p)\in\bar{u}(X)\times\Delta$,

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

as desired.

Conversely, if there exist $\alpha>0$ and $\beta\in\mathbb{R}$ such that $\bar{u}=\alpha u+\beta$ and $\bar{G}(t,p)=\alpha G\left(\alpha^{-1}(t-\beta),p\right)+\beta$ for all $(t,p)\in\bar{u}(X)\times\Delta$, then $\bar{u}:X\to\mathbb{R}$ is affine nonconstant, $\bar{G}:\bar{u}(X)\times\Delta\to(-\infty,\infty]$ belongs to $\mathcal{G}(\bar{u}(X)\times\Delta)$, is linearly continuous, and, for all f and g in $\mathcal{F}$,

$$\begin{aligned}f\succsim g&\iff\inf_{p\in\Delta}G\left(\int u(f)\,dp,p\right)\ge\inf_{p\in\Delta}G\left(\int u(g)\,dp,p\right)\\&\iff\inf_{p\in\Delta}G\left(\int\frac{\bar{u}(f)-\beta}{\alpha}dp,p\right)\ge\inf_{p\in\Delta}G\left(\int\frac{\bar{u}(g)-\beta}{\alpha}dp,p\right)\\&\iff\inf_{p\in\Delta}G\left(\frac{\int\bar{u}(f)\,dp-\beta}{\alpha},p\right)\ge\inf_{p\in\Delta}G\left(\frac{\int\bar{u}(g)\,dp-\beta}{\alpha},p\right)\\&\iff\inf_{p\in\Delta}\left[\alpha G\left(\frac{\int\bar{u}(f)\,dp-\beta}{\alpha},p\right)+\beta\right]\ge\inf_{p\in\Delta}\left[\alpha G\left(\frac{\int\bar{u}(g)\,dp-\beta}{\alpha},p\right)+\beta\right]\\&\iff\inf_{p\in\Delta}\bar{G}\left(\int\bar{u}(f)\,dp,p\right)\ge\inf_{p\in\Delta}\bar{G}\left(\int\bar{u}(g)\,dp,p\right),\end{aligned}$$

and

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

for all $(t,p)\in\bar{u}(X)\times\Delta$. ■

Proof of Proposition 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 $\mathcal{F}$, (7) holds. Then, by A.6, $u(X)=\mathbb{R}$ (see Lemma 69). Set

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

Since $G\in\mathcal{G}(\mathbb{R}\times\Delta)$, then, by Lemma 50, I is finite, (evenly) quasiconcave, monotone, normalized, and (7) amounts to

$$f\succsim g\iff I(u(f))\ge I(u(g)).$$

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 39, and

$$G(t,p) = \sup_{\varphi \in B_0(\Sigma):\langle p,\varphi\rangle \le t} I(\varphi) = \sup_{f \in \mathcal{F}:\langle p,u(f)\rangle \le t} I(u(f)) \qquad \forall (t,p) \in \mathbb{R} \times \Delta.$$

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

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

This proves that that (8) holds, and $G = G^\star$. ■

Proof of Theorem 9. Suppose $\succsim$ satisfies axioms A.1-A.6. By Lemma 67, 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

$$f \succsim g \Longleftrightarrow I(u(f)) \ge I(u(g)).$$

Moreover, by Lemma 69, $u(X) = \mathbb{R}$. Then $B_0(\Sigma, u(X)) = B_0(\Sigma)$. Set

$$G_p(t) = \sup\{I(\varphi) : \varphi \in B_0(\Sigma) \text{ and } \langle p,\varphi\rangle \le t\} \qquad \forall (t,p) \in \mathbb{R}\times\Delta.$$

Theorem 33 guarantees that $I(\varphi) = \min_{p\in\Delta} G_p(\langle p,\varphi\rangle)$ for all $\varphi \in B_0(\Sigma)$. In particular,

$$I(u(f)) = \min_{p\in\Delta} G_p\left(\int u(f)\, dp\right) \qquad \forall f \in \mathcal{F},$$

and (13) holds. Lemmas 28 and 29 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 Proposition 5, $(u, G^\star)$ is an uncertainty averse representation of $\succsim$.

If $\succsim$ also satisfies axiom A.7, that is (i) holds, by Lemma 68, I is uniformly continuous. Then, by Theorem 38, $\mathrm{dom}(G_p) \in \{\emptyset, \mathbb{R}\}$ for all $p \in \Delta$, and $\{G_p\}_{p\in\Delta:\mathrm{dom}(G_p)=\mathbb{R}}$ 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 50,

$$I(\varphi) = \inf_{p\in\Delta} G(\langle p,\varphi\rangle, p) \qquad \forall \varphi \in B_0(\Sigma), \tag{94}$$

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

$$f \succsim g \Longleftrightarrow I(u(f)) \ge I(u(g)). \tag{95}$$

Since G is linearly continuous, I is continuous, thus, by Lemma 67, $\succsim$ satisfies A.1-A.5. Since u is affine, $u(X) = \mathbb{R}$, and u represents $\succsim$ on X, then Lemma 69 guarantees that $\succsim$ satisfies A.6. By Proposition 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 52, $G \in \mathcal{E}(\mathbb{R} \times \Delta)$ implies that I is uniformly continuous, thus, by Lemma 68, $\succsim$ 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 (94).

By Theorem 3, u is cardinally unique, by definition of uncertainty averse representation, for all $(t, p) \in \mathbb{R} \times \Delta$,

$$\begin{aligned} G(t,p) &= \sup_{f\in\mathcal{F}} \left\{ u(x_f) : \int u(f)\,dp \le t \right\} = \sup_{f\in\mathcal{F}} \left\{ I(u(f)) : \int u(f)\,dp \le t \right\} \\ &= \sup\left\{ I(\varphi) : \varphi \in B_0(\Sigma) \text{ and } \langle p, \varphi\rangle \le t \right\} = \sup\left\{ I(\varphi) : \varphi \in B_0(\Sigma) \text{ and } \langle p, \varphi\rangle = t \right\} \\ &= \sup_{f\in\mathcal{F}} \left\{ I(u(f)) : \int u(f)\,dp = t \right\} = \sup_{f\in\mathcal{F}} \left\{ u(x_f) : \int u(f)\,dp = t \right\}, \end{aligned}$$

where the equality in the second line descends from Lemma 31. 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

$$I\left(k1_{E_n} + \psi 1_{E_n^c}\right) = I\left(u(x)\,1_{E_n} + u(f)\,1_{E_n^c}\right) = I\left(u(xE_nf)\right) > I(u(g)) = I(\varphi).$$

By Theorem 53, 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 53, 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

$$I\left(k1_{E_n} + \psi 1_{E_n^c}\right) = I\left(u(x)\,1_{E_n} + u(f)\,1_{E_n^c}\right) = I\left(u(xE_nf)\right)$$

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 Proposition 6. By standard results ([21, Corollary B.3]), (i) implies that $u_1 \approx u_2$. Wlog, $u_1 = u_2 = u$. By (10), 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, $u\left(x_f^2\right) \ge u\left(x_f^1\right)$ for all $f \in \mathcal{F}$. By (8), for all $(t, p) \in u(X) \times \Delta$,

$$G_1(t,p) = \sup_{f\in\mathcal{F}} \left\{ u\left(x_f^1\right) : \int u(f)\,dp \le t \right\} \le \sup_{f\in\mathcal{F}} \left\{ u\left(x_f^2\right) : \int u(f)\,dp \le t \right\} = G_2(t,p),$$

and so $G_1 \le G_2$.

Conversely, assume wlog $u_1 = u_2 = u$. Then, for all $f \in \mathcal{F}$ and $x \in X$,

$$f \succsim_1 x \Longrightarrow \inf_{p\in\Delta} G_1\left(\int u(f)\,dp, p\right) \ge \inf_{p\in\Delta} G_1\left(\int u(x)\,dp, p\right) = u(x),$$

but $G_1 \le G_2$ implies

$$\inf_{p\in\Delta} G_2\left(\int u(f)\,dp, p\right) \ge \inf_{p\in\Delta} G_1\left(\int u(f)\,dp, p\right) \ge u(x) = \inf_{p\in\Delta} G_2\left(\int u(x)\,dp, p\right),$$

which delivers $f \succsim_2 x$. ∎

Proof of Proposition 10. If (u, c) is a variational representation of $\succsim$, it is routine to check that (u, G) is a representation in the sense of (7). Moreover, since $u(X) = \mathbb{R}$, then Proposition 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

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

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 $\operatorname{dom}(c)$ 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,

$$\begin{aligned}c(\alpha p_1+(1-\alpha)p_2)&\leq c(p_2)+t_2-\alpha t_1-(1-\alpha)t_2=c(p_2)+t_2-\alpha t_1-t_2+\alpha t_2\\&=c(p_2)+\alpha(t_2-t_1)=c(p_2)+\alpha(c(p_1)-c(p_2))\\&=\alpha c(p_1)+(1-\alpha)c(p_2),\end{aligned}$$

as wanted. ■

Proof of Theorem 16. Assume (18) holds, i.e.

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

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

$$J(\varphi)=\int_\Delta\phi(\langle p,\varphi\rangle)\,d\mu(p)\in\phi(\mathbb{R})\qquad\forall\varphi\in B_0(\Sigma).\tag{96}$$

By Theorem 54, J is finite, concave, continuous and monotone on X. Therefore the functional

$$I=\phi^{-1}\circ J\tag{97}$$

is well defined, quasiconcave, continuous, monotone, and normalized. Moreover, by (18)

$$f\succsim g\Longleftrightarrow I(u(f))\geq I(u(g)).$$

Thus $\succsim$ satisfies axioms Axiom 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 (66)

$$\sup_{\varphi\in B_0(\Sigma):\langle p,\varphi\rangle\leq t}I(\varphi)=\phi^{-1}\left(\inf\left\{\inf_{k\geq0}\left[tk-\int_\Delta\phi^*\left(k\frac{d\nu}{d\mu}(q)\right)d\mu(q)\right]:\nu\in\Gamma(p)\right\}\right)$$

and

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

by definition of $I_t\left(\cdot\parallel\mu\right)$,

$$\begin{aligned}
&\phi^{-1}\left(\inf\left\{\inf_{k\geq0}\left[tk-\int_{\Delta}\phi^{*}\left(k\frac{d\nu}{d\mu}\left(q\right)\right)d\mu\left(q\right)\right]:\nu\in\Gamma\left(p\right)\right\}\right)\\
&=\inf_{\nu\in\Gamma(p)}\phi^{-1}\left(\inf_{k\geq0}\left[tk-\int_{\Delta}\phi^{*}\left(k\frac{d\nu}{d\mu}\left(q\right)\right)d\mu\left(q\right)\right]\right)=\inf_{\nu\in\Gamma(p)}\left\{I_t\left(\nu\parallel\mu\right)+t\right\}\\
&=t+\inf_{\nu\in\Gamma(p)}I_t\left(\nu\parallel\mu\right),
\end{aligned}$$

and the infimum is attained since $\Gamma(p)$ 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 $(t,p)\in\mathbb{R}\times\Delta$.

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

$$G\left(t,p\right)=t+\min_{\nu\in\Gamma(p)}I_t\left(\nu\parallel\mu\right)\tag{98}$$

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

$$G\left(t,p\right)=\phi^{-1}\left(\inf_{\nu\in\Gamma(p)}\left\{\inf_{k\geq0}\left[kt-\int_{\Delta}\phi^{*}\left(k\frac{d\nu}{d\mu}\right)d\mu\right]\right\}\right),$$

and defining J and I like in (96) and (97), it descends from (66) that

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

Since I is finite, quasiconcave, continuous, monotone, and normalized, by Theorem 33,

$$I\left(\varphi\right)=\inf_{p\in\Delta}G\left(\langle p,\varphi\rangle,p\right)=\min_{p\in\Delta}G\left(\langle p,\varphi\rangle,p\right)\qquad\forall\varphi\in B_0\left(\Sigma\right).\tag{99}$$

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 (98) and (99) imply (23). ■

Proof of Corollary 17. Set $\phi(t)=-e^{-\theta t}$, then $\phi^{-1}(t)=-\theta^{-1}\log(-t)$, and for all $\varphi\in B_0\left(\Sigma\right)$

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

call this functional $I(\varphi)$.

Let $(t,p) \in \mathbb{R} \times \Delta$. By (66)

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

by Proposition 58,

$$\sup_{\varphi \in B_0(\Sigma):\langle p,\varphi\rangle \leq t} I(\varphi) = t + \frac{1}{\theta}\inf\left\{R(\nu \parallel \mu) : \nu \in \Gamma(p)\right\}.$$

Theorem 16 delivers the equivalence between (i) and (ii), while Proposition 10 that between (ii) and (iii).[37] ■

Proof of Theorem 18. By Proposition 57, the functional

$$I(\varphi) = \phi^{-1}\left(\int_\Delta \phi(\langle p,\varphi\rangle)\, d\mu(p)\right) \qquad \forall \varphi \in B_0(\Sigma),$$

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 16, I is well defined, quasiconcave, continuous, monotone, normalized, and

$$f \succsim g \Longleftrightarrow I(u(f)) \geq I(u(g)).$$

But, by definition of variational representation, the functional

$$\bar{I}(\varphi) = \min_{p\in\Delta}(\langle p,\varphi\rangle + c(p)) \qquad \forall \varphi \in B_0(\Sigma),$$

(which is concave, continuous, monotone, normalized, and translation invariant) is such that

$$f \succsim g \Longleftrightarrow \bar{I}(v(f)) \geq \bar{I}(v(g)).$$

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

$$\tilde{I}(\varphi) = \min_{p\in\Delta}(\langle p,\varphi\rangle + \alpha c(p)) \qquad \forall \varphi \in B_0(\Sigma),$$

(which is concave, continuous, monotone, normalized, and translation invariant) is such that

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

By Lemma 67, $I = \tilde{I}$ 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 Axiom A.1-A.6), thus it is a variational preference. ■

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

Proof of Theorem 21. By Theorem 3, G is lower semicontinuous. Set

$$I(\varphi) = \min_{p\in\Delta} G\left(\int \varphi dp, p\right) \qquad \forall \varphi \in B_0(\Sigma).$$

Since (u, G) is a representation, $I : B_0(\Sigma) \to \mathbb{R}$ is normalized and $f \succsim g \iff I(u(f)) \geq I(u(g))$. By Lemmas 67 and 68, $I : B_0(\Sigma) \to \mathbb{R}$ is normalized, monotone, quasiconcave, and uniformly continuous. Moreover, for all $(t, p) \in \mathbb{R} \times \Delta$,

$$G(t,p) = \sup_{f\in\mathcal{F}} \left\{ u(x_f) : \int u(f)\, dp \leq t \right\} = \sup_{f\in\mathcal{F}:\langle p,u(f)\rangle \leq t} I(u(f)) = \sup_{\varphi\in B_0(\Sigma):\langle p,\varphi\rangle\leq t} I(\varphi). \tag{100}$$

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

$$f \sim x_f \Longrightarrow \alpha f + (1-\alpha)\, x_* \sim \alpha x_f + (1-\alpha)\, x_* \tag{101}$$

thus,[38] for all $\phi = u(f) \in B_0(\Sigma)$

$$\begin{aligned} I(\alpha\phi) &= I(\alpha u(f) + (1-\alpha)\, u(x_*)) = I(u(\alpha f + (1-\alpha)\, x_*)) \\ &= I(u(\alpha x_f + (1-\alpha)\, x_*)) = u(\alpha x_f + (1-\alpha)\, x_*) \\ &= \alpha u(x_f) + (1-\alpha)\, u(x_*) = \alpha u(x_f) = \alpha I(u(f)) = \alpha I(\phi). \end{aligned}$$

If $\alpha > 1$ we then have $I(\phi) = I\left(\alpha^{-1}\alpha\phi\right) = \alpha^{-1} I(\alpha\phi)$, and we conclude that $I : B_0(\Sigma) \to \mathbb{R}$ is positively homogeneous. By Theorem 41, this implies (ii).

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

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

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

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

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. ■

[38] Notice that condition (101) is weaker than A.10, and it is sufficient to drive the result.

Proof of Corollary 22. (i) implies (ii). Immediately descends from Theorem 21.

(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

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

By Lemma 44, I is finite, monotone, upper semicontinuous, positively homogeneous, quasiconcave

$$I(\varphi) = \min_{p \in C} \left(\frac{\langle p, \varphi \rangle^+}{c_1(p)} - \frac{\langle p, \varphi \rangle^-}{c_2(p)} \right) = \min_{p \in \Delta} G(\langle p, \varphi \rangle, p) \qquad \forall \varphi \in B_0(\Sigma), \tag{102}$$

and

$$\sup_{\psi \in B_0(\Sigma): \langle p, \psi \rangle \leq t} I(\psi) = G(t,p) \qquad \forall (t,p) \in \mathbb{R} \times \Delta. \tag{103}$$

Moreover, by (103) and Theorem 41, I is monotone, quasiconcave, uniformly continuous, positively homogeneous, and normalized. While, by (28) and (102), for all f and g in $\mathcal{F}$,

$$f \succsim g \Longleftrightarrow I(u(f)) \geq I(u(g)).$$

By Lemmas 67, 68, and 69, $\succsim$ 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 (103) and Theorem 38, (u, G) is a representation of $\succsim$ in the sense of Theorem 9. 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 $\mathcal{F}$,

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

Then u and v represent $\succsim$ on X, therefore there is $\alpha > 0$ such that $v = \alpha u$. Thus

$$f \succsim g \Longleftrightarrow \min_{p \in D} \left(\frac{\left(\int u(f)\, dp\right)^+}{d_1(p)} - \frac{\left(\int u(f)\, dp\right)^-}{d_2(p)} \right) \geq \min_{p \in D} \left(\frac{\left(\int u(g)\, dp\right)^+}{d_1(p)} - \frac{\left(\int u(g)\, dp\right)^-}{d_2(p)} \right). \tag{104}$$

Set

$$J(\varphi) = \min_{p \in D} \left(\frac{\langle p, \varphi \rangle^+}{d_1(p)} - \frac{\langle p, \varphi \rangle^-}{d_2(p)} \right) \qquad \forall \varphi \in B_0(\Sigma).$$

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 (104), for all f and g in $\mathcal{F}$,

$$f \succsim g \Longleftrightarrow J(u(f)) \geq J(u(g)).$$

By Lemma 67, $I = J$ 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, mimeo, 2008.

[8] A. Chateauneuf and J. H. Faro, Ambiguity through confidence functions, mimeo, 2006.

[9] A. Chateauneuf, F. Maccheroni, M. Marinacci and J.-M. Tallon, Monotone continuous multiple priors, *Economic Theory*, 26, 973–982, 2005.

[10] I. Csiszar, I-divergence geometry of probability distributions and minimization problems, *Annals of Probability*, 3, 146–158, 1975.

[11] G. Dal Maso, *An introduction to* Γ*-convergence*, Birkhäuser, Boston, 1993.

[12] G. Debreu, *Theory of value*, Yale University Press, 1959.

[13] B. de Finetti, Sulle stratificazioni convesse, *Annali di Matematica Pura e Applicata*, 30, 173–183, 1949.

[14] B. de Finetti, Sul concetto di media, *Giornale dell'Istituto Italiano degli Attuari*, 2, 369–396, 1931.

[15] 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.

[16] L.G. Epstein and M. Schneider, Recursive multiple-priors, *Journal of Economic Theory*, 113, 1–31, 2003.

[17] H. Ergin and F. Gul, A subjective theory of compound lotteries, mimeo, 2004.

[18] W. Fenchel, A remark on convex sets and polarity, *Meddelanden Lunds Universitets Matematiska Seminarium*, 82–89, 1952.

[19] M. Frittelli, Introduction to a theory of value coherent with the no-arbitrage principle, Finance and Stochastics, 4, 275–297, 2000.

[20] P. Gänssler, Compactness and sequential compactness in spaces of measures, *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 17, 124–146, 1971.

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

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

[23] I. Gilboa, F. Maccheroni, M. Marinacci, and D. Schmeidler, Objective and subjective rationality in a multiple prior model, mimeo, 2008.

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

[25] C-W. Ha, Minimax and fixed point theorems, *Mathematische Annalen*, 248, 73–77, 1980.

[26] E. Hanany and P. Klibanoff, Updating ambiguity averse preferences, mimeo, 2007.

[27] L. P. Hansen, Beliefs, doubts, and learning: valuing macroeconomic risk, *American Economic Review*, 97, 1–30, 2007.

[28] L. P. Hansen and T. Sargent, Robust control and model uncertainty, *American Economic Review*, 91, 60–66, 2001.

[29] 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.

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

[31] M. A. Krasnosel'sii and Y. B. Ruticjii, *Convex Functions and Orlicz Spaces*, Gordon and Breach Science Publishers, New York, 1961.

[32] F. Liese and I. Vajda, *Convex Statistical Distances*, Teubner, Leipzig, 1987.

[33] W. A. J. Luxemburg and A. C. Zaanen, *Riesz spaces*, vol. I, North-Holland, Amsterdam, 1971.

[34] F. Maccheroni, M. Marinacci, and A. Rustichini, Ambiguity aversion, robustness, and the variational representation of preferences, *Econometrica*, 74, 1447–1498, 2006.

[35] F. Maccheroni, M. Marinacci, and A. Rustichini, Dynamic variational preferences, *Journal of Economic Theory*, 128, 4–44, 2006.

[36] M. Marinacci and L. Montrucchio, Introduction to the mathematics of ambiguity, in *Uncertainty in Economic Theory*, (I. Gilboa, ed.), Routledge, New York, 2004.

[37] M. Marinacci and L. Montrucchio, On concavity and supermodularity, *Journal of Mathematical Analysis and Applications*, 344, 642–654, 2008.

[38] R. Nau, Uncertainty aversion with second-order utilities and probabilities, *Management Science*, 52, 136–145, 2004.

[39] J-P. Penot and M. Volle, On quasi-convex duality, *Mathematics of Operations Research,* 15, 597–625, 1990.

[40] L. Rigotti, C. Shannon, and T. Strzalecki, Subjective beliefs and ex-ante trade, *Econometrica*, forthcoming.

[41] R. T. Rockafellar, *Convex analysis*, Princeton University Press, Princeton, 1970.

[42] R. T. Rockafellar, Integrals which are convex functionals, *Pacific Journal of Mathematics*, 24, 525–539, 1968.

[43] R. T. Rockafellar, Integrals which are convex functionals, II, *Pacific Journal of Mathematics*, 39, 439–469, 1971.

[44] W. Schachermayer, Optimal Investment in incomplete markets when wealth may become negative, *Annals of Applied Probability*, 11, 694–734, 2001.

[45] D. Schmeidler, Subjective probability and expected utility without additivity, *Econometrica*, 57, 571–587, 1989.

[46] K. Seo, Ambiguity and second-order belief, *Econometrica*, forthcoming.

[47] M. Siniscalchi, Dynamic choice under ambiguity, mimeo, 2006.

[48] T. Strzalecki, Axiomatic foundations of multiplier preferences, mimeo, 2007

[49] K. Yosida, *Functional analysis*, Springer, New York, 1980.

[50] A. C. Zaanen, *Riesz spaces II*, North-Holland, Amsterdam, 1983.