Ambiguity in the small and in the large*

Paolo Ghirardato[§] Marciano Siniscalchi[¶]

June 2012

Abstract

This paper considers local and global multiple-prior representations of ambiguity for preferences that are (i) monotonic, (ii) Bernoullian, i.e. admit an affine utility representation when restricted to constant acts, and (iii) locally Lipschitz continuous. We do not require either Certainty Independence or Uncertainty Aversion. We show that the set of priors identified by Ghirardato, Maccheroni, and Marinacci (2004)'s 'unambiguous preference' relation can be characterized as a union of Clarke differentials. We then introduce a behavioral notion of 'locally better deviation' at an act, and show that it characterizes the Clarke differential of the preference representation at that act. These results suggest that the priors identified by these preference statements are directly related to (local) optimizing behavior.

*We thank Wolfgang Pesendorfer and two anonymous referees, as well as Simone Cerreia-Vioglio, Theo Diasakos, Fabio Maccheroni, Mark Machina, Pietro Ortoleva, Daniele Pennesi, Tomasz Strzalecki, and audiences at the 2009 RUD and SAET conferences, the Workshop in honor of Daniel Ellsberg (Vienna, May 2010), and at seminars in Rice, Northwestern, Penn, Princeton and Montreal for helpful comments. The usual disclaimer applies. Ghirardato is also grateful to the Italian MIUR for financial support. A version of Theorem 2 in this paper first appeared in the working paper 'A more robust definition of multiple priors' (Ghirardato and Siniscalchi, 2010).

§Department of Economics and Statistics and Collegio Carlo Alberto, Università di Torino; paolo.ghirardato@carloalberto.org

¶Northwestern University; marciano@northwestern.edu

1 Introduction

Several popular models of choice under ambiguity represent preferences over uncertain prospects (acts) via some function of their expected utilities, computed with respect to a distinguished set of probabilities. For instance, the maxmin-expected utility (MEU) model of Gilboa and Schmeidler (1989) ranks acts according to $V(h) = \min_{Q\in D} \mathrm{E}_Q[u \circ h]$, where D is a set of priors over the state space S. For multiplier preferences (Hansen and Sargent, 2001), $V(h) = \min_{Q\in\Delta(S)} \mathrm{E}_Q[u \circ h] + \theta \cdot R(Q\|P)$, where $R(Q\|P)$ is the relative entropy of Q with respect to an approximating model P. In the smooth ambiguity model of Klibanoff, Marinacci, and Mukerji (2005), $V(h) = \int_{\Delta(S)} \phi\left(\mathrm{E}_Q[u \circ h]\right) d\mu(Q)$, where μ is a 'second-order belief' over all priors.

The same preference may admit multiple representations that employ different sets of priors (see Siniscalchi, 2006, for examples). Despite this fact, Ghirardato et al. (2004, GMM henceforth) show that a preference can be associated with a 'canonical' set of priors that is independent of its functional representation. Their identification strategy is as follows. Let $\succcurlyeq$ be the individual's preference; say that act f is 'unambiguously preferred' to act g, written $f \succcurlyeq^* g$, if $f \succcurlyeq g$ and this ranking is preserved across mixtures:

$$\lambda f + (1-\lambda)h \succcurlyeq \lambda g + (1-\lambda)h \qquad \text{for all } \lambda \in (0,1] \text{ and all acts } h. \tag{1}$$

GMM show that, under suitable assumptions, there exist a utility function u and a unique set C of priors such that, for all acts f and g, $f \succcurlyeq^* g$ if and only if $\mathrm{E}_P[u \circ f] \geq \mathrm{E}_P[u \circ g]$ for all $P \in C$ (a representation introduced by Bewley, 2002). Furthermore, under the assumptions in GMM, the representation V of the individual's preferences $\succcurlyeq$ can be written as $V(h) = I(u \circ h)$ for a suitable real functional I; GMM then show that C is the Clarke (1983) differential of I, evaluated at the constant function 0. This characterization makes it practical to compute the set C for many decision models, including MEU, α-MEU, and Choquet-expected utility (Schmeidler, 1989). Notably, these results are *not* restricted to preferences that satisfy Uncertainty Aversion in the sense of Schmeidler (1989).[1]

However, the analysis in GMM has two limitations. First, GMM's differential characterization of the set C depends crucially on the assumption that preferences satisify the 'Certainty

[1] Uncertainty Aversion has been questioned both theoretically and experimentally: cf. Epstein (1999), Ghirardato and Marinacci (2002), Baillon, L'Haridon, and Placido (forthcoming).

Independence' axiom of Gilboa and Schmeidler (1989).[2] This axiom restricts ambiguity attitudes, and rules out several recent models of choice under ambiguity, including multiplier and smooth ambiguity preferences.[3] Second, the results in GMM do not fully reveal the usefulness and economic significance of the set of priors C, beyond the fact that it characterizes the unambiguous preference $\succcurlyeq^*$.

The objective of this paper is to address both limitations. To begin, we do not assume Certainty Independence; as a result, our analysis does not impose any restriction on ambiguity attitudes, and accommodates virtually all classical and recent decision models under ambiguity, including those discussed above or referenced in footnote 3. Our first main result generalizes GMM's differential characterization of the set C: writing the representation of preferences as $V(h) = I(u \circ h)$, we show that, up to convex closure, C is the union of all (suitably normalized) Clarke differentials of the functional I, computed at all interior points rather than just at zero.

The Clarke differential of non-smooth functions plays a similar role in optimization problems as the gradient of smooth functions. In particular, a function attains a local extremum at a point only if its Clarke differential at that point contains the zero vector—an analog of the familiar first-order conditions. Our result then implies that the probabilities in the set C are (up to convex closure) those that *identify candidate solutions to optimization problems.* Example 3 below illustrates this in a canonical portfolio choice application.

Our second main result has no counterpart in GMM, and sheds further light on the role of priors in C in the individual's choices. To illustrate, think of acts f, g as representing the state-contingent consequences of two actions the individual may choose, and act h as the status quo. Then Eq. (1) states that choosing the f action with some probability λ, thereby 'moving' from h toward f in utility terms, is always at least as good as moving toward g, no matter what the initial status-quo point h is and how far one moves away from h. That is, f is a *uniformly bet-*

[2] On the other hand, the existence of a set C of priors that characterizes the unambiguous preference $\succcurlyeq^*$ follows under minimal regularity assumptions: see Sec. 3.

[3] Other models that do not assume Certainty Independence include variational preferences (Maccheroni, Marinacci, and Rustichini, 2006), confidence-function preferences (Chateauneuf and Faro, 2009), uncertainty-averse preferences (Cerreia-Vioglio, Maccheroni, Marinacci, and Montrucchio, 2008), vector expected utility (Siniscalchi, 2009), and mean-dispersion preferences (Grant and Polak, 2011a,b).

ter deviation than g. However, one is typically interested in optimality conditions at a *specific* status-quo point h. With this in mind, we 'localize' Eq. (1); that is, we apply it to a small neighborhood around a single status-quo point h, and only consider small (but discrete) movements away from the status quo. Say that f is a *better deviation than* g *near* h, written $f \succcurlyeq_h^* g$, if

$$\lambda f + (1-\lambda)h' \succcurlyeq \lambda g + (1-\lambda)h' \qquad \text{for all } \lambda \text{ small and all acts } h' \text{ near } h \tag{2}$$

(see Sec. 4.3 for details). This definition is naturally related to optimizing behavior at a point h: it identifies small profitable and unprofitable deviations away from the status quo. Our second result shows that the relation $\succcurlyeq_h^*$ characterizes the normalized Clarke differential $C(h)$ of the functional I at the point h.[4] In our view, this result illustrates the connection between priors and optimizing behavior in a clearer and more direct way than GMM's global result (and our generalization thereof). That said, the local and global results are closely related: $f \succcurlyeq^* g$ if and only if $f \succcurlyeq_h^* g$ for all acts h, and C is the union of all sets $C(h)$.

Finally, a caveat. GMM suggest that the set C may represent 'ambiguous beliefs' or 'perceived ambiguity.' However, they also discuss (GMM, p. 137) potential difficulties with this interpretation; in particular, C may incorporate aspects of ambiguity attitude. We prefer to emphasize the connection between the priors in the set C and optimizing *behavior*, and do not take a stand as to whether such priors reflect beliefs, ambiguity attitudes, or both.

1.1 Intuition for the results and examples

For simplicity, let the state space be $S = \{s_1, s_2\}$ and assume linear utility. To make the intuition as sharp as possible, we assume that I is continuously differentiable, so its Clarke differential at a point h is the gradient $\nabla I(h)$, and, importantly, the map $h \mapsto \nabla I(h)$ is continuous.

Under these assumptions, C is the convex closure of the set of all the probabilities $\frac{\nabla I(k)}{\nabla I(k)\cdot[1,1]}$ for all $k \in \mathbb{R}^2$. To see that, for every $f, g \in \mathbb{R}^2$, $P \cdot f \geq P \cdot g$ for all $P \in C$ implies Eq. (1), fix $h \in \mathbb{R}^2$ and $\lambda \in (0,1]$. By assumption, $\nabla I(k) \cdot (f-g) \geq 0$ for all $k \in \mathbb{R}^2$: then, by the mean value theorem, there is a point k^* in the segment joining $\lambda f + (1-\lambda)h$ and $\lambda g + (1-\lambda)h$ such that

[4] $\succcurlyeq_h^*$ is not a Bewley preference, and its connection with $C(h)$ is more subtle than the relationship between $\succcurlyeq^*$ and C: see Sec. 4.3 for details.

$I(\lambda f+(1-\lambda)h)-I(\lambda g+(1-\lambda)h)=\nabla I(k^*)\cdot[\lambda f+(1-\lambda)h-\lambda g+(1-\lambda)h]=\nabla I(k^*)\cdot\lambda(f-g)\geq 0$, so Eq. (1) holds. This argument generalizes to the non-smooth case.

For the converse implication, suppose that $P^*\cdot f < P^*\cdot g$ for some $P^*\in C$, and hence $\nabla I(k^*)\cdot(f-g)<0$ for some k^*; then, since $\nabla I(k)$ is continuous in k, $\nabla I(k)\cdot(f-g)<0$ for all k in some neighborhood N of k^*. But then we can let $h=k^*$ and choose λ sufficiently small so that the segment joining $\lambda f+(1-\lambda)k^*$ and $\lambda g+(1-\lambda)k^*$ lies entirely in N; thus, the mean value theorem implies that $I(\lambda f+(1-\lambda)k^*)-I(\lambda g+(1-\lambda)k^*)<0$, so Eq. (1) does not hold. This argument relies crucially on the fact that there is a unique gradient at every point k, and that the gradient is continuous in k. Both properties fail in the non-smooth case, so our proof of Theorem 2 takes a different approach.

Turn now to our local characterization result: $C(h)=\{P(h)\}$, where $P(h)=\frac{\nabla I(h)}{\nabla I(h)\cdot[1,1]}$. Assume that Eq. (2) holds, and in particular consider the sequence $h^n=\frac{1}{1-\lambda^n}h$. Then it is easy to see that, for all n large, $\frac{I(\lambda f+h)-I(h)}{\lambda^n}\geq\frac{I(\lambda g+h)-I(h)}{\lambda^n}$; since I is differentiable, this implies that $\nabla I(h)\cdot f\geq\nabla I(h)\cdot g$. Thus, $f\succcurlyeq^*_h g$ implies that $P(h)\cdot f\geq P(h)\cdot g$. Here, differentiability allows us to focus on a specific sequence (h^n), and directly link Eq. (2) to a property of the unique differential of I at h. The non-smooth case again requires a different approach.

The converse implication is more delicate, *even in the smooth case*. By differentiability, if $\nabla I(h)\cdot f>\nabla I(h)\cdot g$, then $\frac{I(\lambda f+h)-I(h)}{\lambda^n}>\frac{I(\lambda g+h)-I(h)}{\lambda^n}$ for n large, so Eq. (2) holds for the sequence $h^n=\frac{1}{1-\lambda^n}h$ considered above. To to extend this conclusion to other sequences, one needs to invoke the fact that a continuously differentiable function is 'strictly differentiable' (Clarke, 1983, Prop. 2.2.1). But, if $\nabla I(h)\cdot f=\nabla I(h)\cdot g$, this argument clearly does not apply. Example 4 in Sec. 4.3 illustrates further subtleties. Theorems 6 and 7 circumvent these issues.

Example 1 (Non-smooth preferences) Example 17 in GMM characterizes the set C for Choquet preferences on a finite state space $S=\{s_1,\dots,s_n\}$. We briefly discuss the characterization of the local priors $C(h)$. For any permutation σ of $\{1,\dots,n\}$, a Choquet preference with capacity v admits an EU representation on the set $\mathscr{F}_\sigma$ of acts h such that $h(s_{\sigma(1)})\succcurlyeq h(s_{\sigma(2)})\succcurlyeq\ldots\succcurlyeq h(s_{\sigma(n)})$, with prior P_σ given by $P_\sigma(s_{\sigma(i)})=v(\{s_{\sigma(1)},\dots,s_{\sigma(i)}\})-v(\{s_{\sigma(1)},\dots,s_{\sigma(i-1)}\})$. Fix an act h that belongs only to $\mathscr{F}_\sigma$; preferences are effectively EU in a 'neighborhood' of h, so $C(h)=\{P_\sigma\}$. If instead h belongs to $\mathscr{F}_{\sigma_1},\dots,\mathscr{F}_{\sigma_k}$, then, by Theorem 2.5.1 in Clarke (1983), $C(h)$ is the convex hull of

$\{P_{\sigma_1}, \ldots, P_{\sigma_k}\}$. This result extends to piecewise linear preferences (defined in GMM, §5.2).

Example 2 (Local vs. global priors) Let $S = \{s_1, s_2\}$, $X = \mathbb{R}_+$, and the risk-neutral preference represented by $I(h) = \max\left(\frac{1}{2}h(s_1) + \frac{1}{2}h(s_2), \epsilon + \min(h(s_1), h(s_2))\right)$, for some small $\epsilon > 0$.[5] For acts h such that $|h(s_1) - h(s_2)| \geq 2\epsilon$, preferences are consistent with EU, with a uniform prior P on S; if ϵ is small, this is the case for 'most' acts. However, for acts close to the diagonal, this preference behaves like MEU, with set of priors $\Delta(S)$.

Our generalization of GMM's result implies that the preference $\succcurlyeq^*$ defined in Eq. (1) is represented by $C = \Delta(S)$, despite the fact that I is consistent with EU for 'most' acts; we view this as a stark demonstration of the global nature of GMM's approach. By way of contrast, $C(h) = \{P\}$ if $|h(s_1) - h(s_2)| > \epsilon$, and $C(h) = \Delta(S)$ if $|h(s_1) - h(s_2)| < 2\epsilon$, and our second main result implies that Eq. (2) correctly reflects the local behavior of this preference. ■

Example 3 (based on Dow and da Costa Werlang (1992)) An investor with wealth W and preferences characterized by the functional I and the utility $u : X \to \mathbb{R}$ (where $X \subset \mathbb{R}$) considers buying or selling an asset with uncertain returns $R : \Omega \to \mathbb{R}$ on the finite state space S, at a price p. Thus, the agent's utility if she buys $t \in \mathbb{R}$ units of the asset is $I(u(W + t[R - p]))$. Dow and da Costa Werlang (1992) assume that I is an uncertainty-averse (i.e. concave) Choquet functional (Schmeidler, 1989) and that u is strictly increasing and continuously differentiable; they show that the agent will optimally choose $t = 0$ (i.e. 'no trade') iff $I(R) \leq p \leq -I(-R)$.

We now generalize this result. Assume that I is locally Lipschitz continuous; then (see Clarke, 1983, Prop. 2.3.2), a *necessary* condition for no trade to be optimal is that 0 be an element of the Clarke differential of the real function $t \mapsto I(u(W + t[R - p]))$ at $t = 0$. By the chain rule for non-smooth functions (see Clarke, 1983, Prop. 2.3.9), this translates to: $\mathrm{E}_Q[u'(W)(R - p)] = 0$ for some $Q \in \partial I(u(W))$, the Clarke differential of I at $u(W)$. This generalizes the familiar first-order condition with EU preferences. Moreover, since W is constant and $u'(W) > 0$, we obtain

$$\min_{P \in C(1_S W)} \mathrm{E}_P[R] \leq p \leq \max_{P \in C(1_S W)} \mathrm{E}_P[R],$$

where $C(1_S W)$ is the normalized Clarke differential characterized by Eq. (2), at $h = 1_S W$.

[5] We thank an anonymous referee for suggesting this example.

If, furthermore, the functional I and the function u are concave, this condition is also *sufficient*. This generalizes the result of Dow and Werlang to a broad class of uncertainty-averse preferences. Indeed, the above condition is also sufficient as long as the composite map $t \mapsto I(u(W + t[R - p])$ is concave, even though I is not. For instance (cf. Heath and Tversky, 1991), the investor may be uncertainty-averse with respect to R, yet feel 'competent' enough to evaluate other prospects in a manner consistent with uncertainty appeal.

Finally, if I is an uncertainty-averse Choquet functional, by Corollary 5 $C(1_S W) = C(0_S) = C$, the GMM set of priors. But, since uncertainty-averse Choquet preferences are MEU, $I(u \circ h) = \min_{P \in C} \mathrm{E}_P[u \circ h]$. This yields Dow and Werlang's original result as a special case. ■

1.2 Related literature

As noted above, GMM is the starting point of our work. The discussion of Corollaries 3–5 in Section 4.2 explains how our result specializes to GMM's under Certainty Independence. Nehring (2002) also identifies the set C from behavior; our paper thus also extends his results.

Gilboa, Maccheroni, Marinacci, and Schmeidler (2010) consider a DM who is endowed with a possibly incomplete preference over acts reflecting 'objective' information, and a complete preference reflecting her actual behavior. The objective preference has a Bewley-style representation via a set C of priors. Thus, while there are natural similarities, our objectives are clearly different. We do not posit the existence of objective information. Moreover, our main contribution is the operational characterization of the sets C and $C(h)$.

Siniscalchi (2006) proposes a related notion of 'plausible priors.' The main difference with the present paper, and with the GMM approach, is the fact that plausible priors are identified individually, rather than as elements of a set. This requires restrictions on preferences that we do not need (in addition to Certainty Independence).

Klibanoff, Mukerji, and Seo (2011, KMS henceforth) consider infinite repetitions of an experiment with outcomes in some set S, and impose a 'symmetry' requirement on preferences. They show that, in this setting, $C = \left\{ \int \ell^\infty dm(\ell) : m \in M \right\}$, where ℓ^∞ denotes the i.i.d. product of $\ell \in \Delta(S)$, and $M \subset \Delta(\Delta(S))$. KMS propose a 'relevance' condition that identifies measures in the support of some $m \in M$. This approach differs substantially from GMM's identification strategy. For instance, consider an EU preference with a prior P that, by the symmetry requirement, satisfies $P = \int \ell^\infty dm$ for some $m \in \Delta(\Delta(S))$. KMS deem 'relevant' all measures in the support of m, whereas GMM (and we) find that $C = \{P\}$.

None of the above papers provide a counterpart to our local characterization result.

Rigotti, Shannon, and Strzalecki (2008) propose different, equivalent notions of 'belief at an act h' in a setting with monetary outcomes and preferences represented by a quasiconcave function V, and use them to analyze efficiency and trade in a competitive environment. When $V(h) = I(u(h))$ and I and u are suitably regular,[6] their 'beliefs at h' can be computed from the set $C(h)$ that we characterize and the derivative of u, via an appropriate chain rule (cf. Example 3). Thus, up to marginal utilities, our Eq. (2) provides a complementary behavioral interpretation of Rigotti et al. (2008)'s beliefs at h, and relates these to the GMM set of priors C. On the other hand, our results do not require quasiconcavity.

Finally, Machina (2005) defines 'event derivatives' of a representation $V(\cdot)$, a subjective counterpart to derivatives with respect to lotteries in Machina (1982). A representation is 'event-smooth' if it admits suitably regular event derivatives. Machina shows how to generalize EU-based characterizations of e.g. likelihood rankings or comparative risk aversion to event-smooth representations of preferences; however, his paper does not provide a preference foundation for event smoothness. Instead, our paper focuses on the behavioral properties that characterize the normalized Clarke differential $C(h)$ at an act h. At a formal level, we consider Clarke derivatives with respect to outcomes, rather than events, and do not assume smoothness.

2 Notation and preliminaries

We consider a state space S, endowed with a sigma-algebra Σ. The notation $B_0(\Sigma, \Gamma)$ indicates the set of simple Σ–measurable real functions on S with values in the interval[7] $\Gamma \subset \mathbb{R}$, endowed with the topology induced by the supremum norm; for simplicity, write $B_0(\Sigma, \mathbb{R})$ as $B_0(\Sigma)$. Recall that, since Σ is a sigma-algebra, $B(\Sigma)$ is the closure of $B_0(\Sigma)$, and it is a Banach space.

[6] In particular, if I is locally Lipschitz and nice in the sense of Sec. 4.1, and u is differentiable.

[7] That is, $\Gamma \subset \mathbb{R}$ is one of $[\alpha, \beta]$, $[\alpha, \beta)$, $(\alpha, \beta]$, or (α, β), where $\alpha = -\infty$ and $\beta = \infty$ are allowed where applicable.

The set of finitely additive probabilities on Σ is denoted $ba_1(\Sigma)$. $ba_1(\Sigma)$ is endowed with the (relative) weak* topology; i.e., $\sigma(ba(\Sigma), B_0(\Sigma))$ (equivalently, $\sigma(ba(\Sigma), B(\Sigma))$). We identify elements of $ba(\Sigma)$ and the linear functionals they identify; if $a \in B(\Sigma)$ and $Q \in ba(\Sigma)$, $Q(a) = \int a\, dQ$.

If B is one of $B_0(\Sigma, \Gamma)$ for some interval Γ or $B(\Sigma)$, a functional $I : B \to \mathbb{R}$ is: **monotonic** if $I(a) \geq I(b)$ for all $a \geq b$; **continuous** if it is sup-norm continuous; **isotone** if, for all $\alpha, \beta \in \Gamma$, $I(\alpha 1_S) \geq I(\beta 1_S)$ if and only if $\alpha \geq \beta$; **normalized** if $I(\alpha 1_S) = \alpha$ for all $\alpha \in \Gamma$; **constant-additive** if $I(a + \alpha 1_S) = I(a) + \alpha$ for all $a \in B$ and $\alpha \in \mathbb{R}$ such that $a + \alpha 1_S \in B$; **positively homogeneous** if $I(\alpha a) = \alpha I(a)$ for all $a \in B$ and $\alpha \in \mathbb{R}_+$ such that $\alpha a \in B$; and **constant-linear** if it is constant-additive and positively homogeneous.

Finally, fix a convex subset X of a vector space. (Simple) acts are Σ-measurable functions $f : S \to X$ such that $f(S) = \{f(s) : s \in S\}$ is finite; the set of all (simple) acts is denoted by $\mathscr{F}$. We define mixtures of acts pointwise: for any $\alpha \in [0,1]$, $\alpha f + (1-\alpha) g$ is the act that delivers the prize $\alpha f(s) + (1-\alpha) g(s)$ in state s. Given a preference $\succcurlyeq$ on $\mathscr{F}$, we say that an act $h \in \mathscr{F}$ is *interior* if there exist prizes $x, y \in X$ such that $x \succ h(s) \succ y$ for all $s \in S$, and we denote the set of interior acts by $\mathscr{F}^{\text{int}}$. (The dependence of $\mathscr{F}^{\text{int}}$ on $\succcurlyeq$, while not made explicit, should be kept in mind.)

3 Preferences

The main object of study is a binary relation $\succcurlyeq$ on $\mathscr{F}$. As usual, $\succ$ (resp. $\sim$) denotes the asymmetric (resp. symmetric) component of $\succcurlyeq$. With a small abuse of notation, we denote with the same symbol the prize x and the constant act that delivers x for all s. We assume throughout that the preference $\succcurlyeq$ admits a numerical representation that satisfies a regularity property:

Definition 1 A preference relation $\succcurlyeq$ is (non-trivial) **monotonic, Bernoullian, and Locally Lipschitzian** (henceforth **MBL**) if there exists a non-costant, affine function $u : X \to \mathbb{R}$ and a monotonic, isotone functional $I : B_0(\Sigma, u(X)) \to \mathbb{R}$ that is locally Lipschitz in the interior of its domain, and such that, for all $f, g \in \mathscr{F}$.

$$f \succcurlyeq g \Longleftrightarrow I(u \circ f) \geq I(u \circ g). \tag{3}$$

MBL preferences admit *certainty equivalents*: for any $f \in \mathscr{F}$, there is $x_f \in C$ such that $x_f \sim f$.

Most preference models considered in the classic and recent literature on ambiguity belong to this class. Virtually all have monotonic, isotone, Bernoullian, and continuous representations; Cerreia-Vioglio, Ghirardato, Maccheroni, Marinacci, and Siniscalchi (2011) provide an axiomatization of preferences satisfying these assumptions. Furthermore, if the representing functional is also *constant-additive*, it is globally Lipschitz; this applies to the preferences considered by GMM (including MEU, α-MEU and CEU), as well as multiplier, variational, VEU and mean-dispersion preferences (Grant and Polak, 2011a). Alternatively, if the representing functional is *concave or convex*, then it is locally Lipschitz on the interior of its domain, by a classic result of Roberts and Varberg (1974): this includes smooth uncertainty-averse preferences and the confidence-function preferences studied by Chateauneuf and Faro (2009). Also, if I is *continuously Frechet differentiable*, then again it is locally Lipschitz (see Clarke, 1983, Prop. 2.2.1 and Corollary).

In addition, Online Appendix A introduces a novel axiom that is equivalent to the existence of a locally Lipschitz, normalized representation for monotonic, Bernoullian and continuous preferences. This enables us to apply our results below even to preferences that do not fall into any of the above categories: for instance, uncertainty-averse preferences that are not concavifiable, or the generalized mean-dispersion preferences of Grant and Polak (2011b).

Though MBL preferences are more general than those considered in GMM, these authors' notion of 'unambiguous preference' still identifies a unique set of priors via a Bewley-like representation. The proof is a straightforward adaptation of GMM's, and hence omitted. As we argued in the Introduction, it is useful to interpret GMM's definition as stating that an act f is a better deviation than another act g regardless of what the 'status-quo' h is, and regardless of how far one moves away from h (i.e. how much weight one places on f or, respectively, g).

Definition 2 Let $f, g \in \mathscr{F}$. We say that f is a **uniformly (weakly) better deviation than** g, denoted by $f \succcurlyeq^* g$, if and only if, for each $h \in \mathscr{F}$ and each $\lambda \in (0,1]$, $\lambda f + (1-\lambda)h \succcurlyeq \lambda g + (1-\lambda)h$.

Proposition 1 (GMM, Propositions 4 and 5) *Let $\succcurlyeq$ be an MBL preference. Then, there exists a non-empty, unique, convex and closed set $C \subset ba_1(\Sigma)$ such that for each $f, g \in \mathscr{F}$,*

$$f \succcurlyeq^* g \quad \Longleftrightarrow \quad \int u \circ f \, dP \geq \int u \circ g \, dP \quad \text{for all } P \in C, \tag{4}$$

where u is the function in Def. 1. is independent of the choice of normalization of u.

The set C in Proposition 1 is the set of **relevant priors** for the preference $\succcurlyeq$. The following terminology is convenient: a binary relation $\widehat{\succcurlyeq}$ on $\mathscr{F}$ admits a **Bewley representation** (and hence is a **Bewley preference**) if there are an affine function $v : X \to \mathbb{R}$ and a set $D \subset ba_1(\Sigma)$ such that $f \widehat{\succcurlyeq} g$ if and only if $P(v \circ f) \geq P(v \circ g)$ for all $P \in D$.[8] Thus, Proposition 1 states that $\succcurlyeq^*$ is a Bewley preference represented by u and C.

4 Relevant priors: characterizations

4.1 Clarke differentials

Definition 3 (Clarke 1983, Sec. 2.1; Lebourg 1979, Sec. 1) Let B denote either $B_0(\Sigma)$ or $B(\Sigma)$. Consider a locally Lipschitz functional $I : U \to \mathbb{R}$, where $U \subset B$ is open. For every $c \in U$ and $a \in B(\Sigma)$, the **Clarke (upper) derivative** of I in c in the direction a is

$$I^{\circ}(c;a) = \limsup_{b \to c, t \downarrow 0} \frac{I(b+ta) - I(b)}{t}.$$

The **Clarke (sub)differential** of I at c is the set

$$\partial I(c) = \{Q \in ba(\Sigma) : Q(a) \leq I^{\circ}(c;a), \ \forall a \in B\}.$$

It is important to point out that, like the usual notion of gradient, the definition of Clarke differential is seldom used directly (although we do so in proving the results in this section). It is useful chiefly because of its convenient calculus properties (see e.g. Clarke, 1983).

Consider an MBL preference with representation (I, u). Given an interior act h, the functionals in $\partial I(u \circ h)$ are linear, but in general not normalized. For consistency with the GMM approach, we normalize the elements of $\partial I(u \circ h)$ to obtain:

$$C(h) = \left\{ \frac{Q}{Q(S)} : Q \in \partial I(u \circ h), \ Q(S) > 0 \right\}. \tag{5}$$

Given $C(h) \neq \emptyset$ and u, we can define a Bewley preference $\succcurlyeq_{C(h)}$ on $\mathscr{F}$ as follows:

$$f \succcurlyeq_{C(h)} g \qquad \Leftrightarrow \qquad P(u \circ f) \geq P(u \circ h) \quad \forall P \in C(h). \tag{6}$$

[8] Clearly, a set $D \subset ba_1(\Sigma)$ and its convex closure $\overline{\text{co}}\ D$ induce the same Bewley preference. Prop. A.2 in GMM characterizes Bewley preferences.

Say that the functional I is **nice at** $c \in \text{int } B_0(\Sigma, \Gamma)$ if the zero measure $Q_0 \in ba(\Sigma)$ is not an element of $\partial I(c)$. This condition strengthens monotonicity and loosely speaking, requires that preferences remain non-trivial in arbitrarily small neighborhoods of an act. It plays a role in our local results (Propositions 6 and 7), though not in our global result (Theorem 2). All preferences considered by GMM and, more generally, all MBL preferences represented by a *constant-additive* functional I, are everywhere nice; the same is true for *concave* preferences: see Online Appendix B. To cover the remaining cases, Online Appendix B also provides an axiom for arbitrary MBL preferences that ensures the existence of a nice representation.

4.2 Global Characterization

We are ready to state our first main result.

Theorem 2 *For any MBL preference $\succsim$ with representation (I, u) and relevant priors C,*

$$C = \overline{\text{co}} \left(\bigcup_{h \in \mathscr{F}^{\text{int}}} C(h) \right).$$

Proof: See Appendix A.2. ■

Thus, up to convex closure, the set C can be computed by considering the normalized Clarke differentials $C(h)$ for all interior acts h, then taking the union of such objects. Equivalently, $f \succsim^* g$ if and only if $P(u \circ f) \geq P(u \circ g)$ for all $P \in C(h)$ and $h \in \mathscr{F}^{\text{int}}$.

We now review specific independence properties of the preference $\succsim$ that have been analyzed in the literature. This will also clarify the relationship between Theorem 2 and its counterpart in GMM. First, if preferences satisfy the 'Weak Certainty Independence' axiom of Maccheroni et al. (2006), the functional I is constant-additive, and all elements $Q \in \partial I(e)$ satisfy $Q(S) = 1$ (see part 2 of Prop. A.3 in GMM). Thus, $C(h) = \partial I(u \circ h)$, and we obtain

Corollary 3 *If I is normalized and constant-additive, then $C = \overline{\text{co}} \left(\bigcup_{h \in \mathscr{F}^{\text{int}}} \partial I(u \circ h) \right)$.*

If instead an MBL preference satisfies the 'Homotheticity' axiom of Cerreia-Vioglio et al. (2008), I is positively homogeneous. If I is normalized, and there is a prize $z \in X$ with $u(z) = 0 \in \text{int } u(X)$, then $\partial I(u \circ h) \subset \partial I(0)$ for all $h \in \mathscr{F}^{\text{int}}$ (cf. part 1 of Prop. A.3 in GMM). We obtain

Corollary 4 *If I is normalized and positively homogeneous, and z is as above, then $C = \overline{\text{co}}\, C(z)$.*

Finally, GMM consider preferences that satisfy 'Certainty Independence,' and hence admit a representation with I normalized and constant-linear. With z as above, we obtain

Corollary 5 (GMM, Theorem 14) *If I is constant-linear, then $C = \partial I(0) = C(z)$.*

Notice that, if I is constant-linear, the Clarke upper derivative at 0 in the direction $a \in B$ takes a particularly simple form (cf. GMM, Prop. A.3), which GMM exploit in their proofs:

$$I^{\circ}(0; a) = \sup_{b \in B} I(b + a) - I(b).$$

4.3 Local Characterization

We turn to the behavioral characterization of 'locally relevant' priors. Recalling the discussion in the Introduction, the definition of locally better deviation concerns the behavior of $\succcurlyeq$ 'near' an interior act h.[9] Thus, its formal statement requires a notion of convergence for acts. We say that a sequence $(f^n) \subset \mathscr{F}$ **converges to** an act $f \in \mathscr{F}$, written $f^n \to f$, iff, for all prizes $x, y \in X$ with $x \succ y$, there exists K such that $k \geq K$ implies

$$\forall s \in S, \qquad \frac{1}{2} f(s) + \frac{1}{2} y \prec \frac{1}{2} f^k(s) + \frac{1}{2} x \quad \text{and} \quad \frac{1}{2} f^k(s) + \frac{1}{2} y \prec \frac{1}{2} f(s) + \frac{1}{2} x.$$

This corresponds to uniform convergence in $B_0(\Sigma, u(X))$.

Intuitively, we then apply Def. 2 to a neighborhood of an interior act h: we consider mixtures of the acts f and g with an act 'near h,' assigning 'most' of the weight to the latter.

Definition 4 For any triple of acts $f, g, h \in \mathscr{F}$, say that f **is a (weakly) better deviation than** g **near** h, written $f \succcurlyeq_h^* g$, if, for every $(\lambda^n)_{n \geq 0} \subset [0, 1]$ and $(h^n)_{n \geq 0} \subset \mathscr{F}$ such that $\lambda^n \downarrow 0$ and $h^n \to h$,

$$\lambda^n f + (1 - \lambda^n) h^n \succcurlyeq \lambda^n g + (1 - \lambda^n) h^n \qquad \text{eventually.} \tag{7}$$

Unlike $\succcurlyeq^*$, the relation $\succcurlyeq_h^*$ is not always a Bewley preference, because it may fail continuity (Online Appendix E indicates the properties it does satisfy). Our first main result in this section shows that it nonetheless uniquely identifies the set $C(h)$.

Theorem 6 *For any MBL preference $\succcurlyeq$ with representation (I, u), and any interior act $h \in \mathscr{F}$:*

[9] Clearly, given a representation (I, u) of $\succcurlyeq$, $h \in \mathscr{F}^{\text{int}}$ if and only if the function $u \circ h$ is in the interior of $B_0(\Sigma, u(X))$.

1. *for all $f, g \in \mathscr{F}$, $f \succsim_h^* g$ implies that $P(u \circ f) \geq P(u \circ g)$ for all $P \in C(h)$;*

2. *if I is nice at $u \circ h$, then the preference $\succsim_{C(h)}$ is the unique minimal Bewley preference that extends $\succsim_h^*$ (i.e., the intersection of all Bewley preferences that contain $\succsim_h^*$).*

Proof: See Appendix A.1. ■

Thus, $f \succsim_h^* g$ always implies $f \succsim_{C(h)} g$; moreover, if I is nice at $u \circ h$, $\succsim_h^*$ fully identifies the set $C(h)$. There is however a different, more direct, way to identify $C(h)$ from $\succsim_h^*$. The following example illustrates this idea, as well as the role of the niceness assumption.

Example 4 Let $S = \{s_1, s_2\}$, $X = \mathbb{R}$, and let u be the identity. Thus, $\mathscr{F} = B_0(2^S, u(X)) = \mathbb{R}^2$, and we identify acts h with vectors $[h_1, h_2] \in \mathbb{R}^2$. Fix $p \in \left(\frac{1}{2}, 1\right)$, let $P^1 = [p, 1-p]$ and $P^2 = [1-p, p]$, and consider the smooth ambiguity preference represented by u and by the strictly increasing and continuously differentiable (hence, locally Lipschitz) function $I : \mathbb{R}^2 \to \mathbb{R}$ defined by $I(h) = \sum_{i=1,2}(P^i \cdot h)^3$. For any h, the Clarke differential $\partial I(h)$ coincides with the gradient $\nabla I(h) = 3(P^1 \cdot h)^2 P^1 + 3(P^2 \cdot h)^2 P^2$. Therefore, for $h \neq 0$, $C(h) = \left\{\frac{1}{3}\nabla I(h)\right\}$, while $C(0) = \emptyset$ since $\nabla I(0) = 0$.

As expected, $f \succsim_h^* g$ implies $f \succsim_{C(h)} g$. As to the opposite implication, we argue in Online Appendix C that, if $f \succ_{C(h)} g$ —i.e., $\nabla I(h) \cdot f > \nabla I(h) \cdot g$— then Eq. (7) eventually holds for all sequences $\lambda^k \downarrow 0$ and $h^n \to h$. We now show that, if $f \sim_{C(h)} g$ or $C(h) = \emptyset$, Eq. (7) does not necessarily hold. Again, details and proofs of all claims below are found in Online Appendix C.

Let $f = [1, 0]$ and $g = [0, 1]$. Fix $h \in \mathbb{R}^2$ arbitrarily, and define sequences $(\lambda^n), (h^n), (k^n)$ by

$$\lambda^n = \frac{1}{n}, \quad h^n = \frac{1}{1-\lambda^n}(\lambda^n[2,1] + h), \quad k^n = \frac{1}{1-\lambda^n}(\lambda^n([1,2] + h).$$

Notice that as $n \to \infty$, $\lambda^n \downarrow 0$, $h^n \to h$ and $k^n \to h$. Then $f \succsim_h^* g$ requires that, for n large,

$$\lambda^n f + (1-\lambda^n)h^n \succsim \lambda^n g + (1-\lambda^n)h^n \quad \text{and} \quad \lambda^n f + (1-\lambda^n)k^n \succsim \lambda^n g + (1-\lambda^n)k^n. \tag{8}$$

We focus on three cases (we omit the other cases for brevity).

<u>Case 1: $h_1 > h_2 \geq 0$</u>. In this case, $f \succ_{C(h)} g$. As noted above, this implies $f \succsim_h^* g$.

<u>Case 2: $h_1 = h_2 = \gamma > 0$</u>. In this case $\nabla I(h) \cdot f = \nabla I(h) \cdot g$, i.e. $f \sim_{C(h)} g$. Then, for n large, the second preference in Eq. (8) is violated, so it is *not* the case that $f \succsim_h^* g$. However, clearly $\nabla I(h) \cdot (f + \epsilon) > \nabla I(h) \cdot (g - \epsilon)$ for any $\epsilon > 0$. As noted above, this implies that $f + \epsilon \succsim_h^* g - \epsilon$.

Case 3: $h_1 = h_2 = 0$. Since $\nabla I(0) = 0$, I is not nice at $h = 0$. Then Eq. (8) does not hold, and continues to be violated if f and g are replaced with with $f + \epsilon$ and $g - \epsilon$, for $\epsilon > 0$ sufficiently small. Thus, neither $f \succcurlyeq^*_h g$ nor $f + \epsilon \succcurlyeq^*_h g - \epsilon$ hold. ■

Example 4 suggests that, while $f \succcurlyeq_{C(h)} g$ may not imply $f \succcurlyeq^*_h g$, it may still imply that $f' \succcurlyeq^*_h g'$ for all f', g' with $f'(s) \succ f(s)$ and $g(s) \succ g'(s)$ for all $s \in S$; we call such a pair of acts (f', g') a **spread** of (f, g). The following result confirms that this is indeed the case, and provides a direct characterization of $C(h)$ in terms of $\succcurlyeq^*_h$.

Theorem 7 *Consider an MBL preference $\succcurlyeq$ and a representation (I, u). Fix an interior act $h \in \mathscr{F}$ and assume that I is nice at $u \circ h$. Then $C(h)$ is the only weak*-closed, convex set $D \subset ba_1(\Sigma)$ for which the following statements are equivalent for every pair (f, g) of interior acts:*

(1) $f' \succcurlyeq^*_h g'$ for all spreads (f', g') of (f, g).

(2) $P(u \circ f) \geq P(u \circ g)$ for all $P \in D$

Proof: See Appendix A.1. ■

Case (3) in the Example shows that niceness is required in Theorems 6 and 7.[10]

Finally, as noted in the Introduction, there is a tight connection between the global preference $\succcurlyeq^*$ and the local preferences $\succcurlyeq^*_h$:

Corollary 8 *For all $f, g \in \mathscr{F}$: $f \succcurlyeq^* g$ if and only of $f \succcurlyeq^*_h g$ for all interior $h \in \mathscr{F}$.*

Proof: See Appendix A.2. ■

5 Extensions

All the results in this paper apply *verbatim* if preferences are defined on the set of bounded (rather than simple) acts, as defined e.g. in Gilboa and Schmeidler (1989).

[10] In Example 4, $\partial I(0)$ contains *only* the zero vector. However, we show in Online Appendix C how to modify preferences so that $\partial I(0)$ contains vectors other than 0, without changing the conclusions of the example.

Theorem 2 can be generalized to preferences that are continuous but possibly not locally Lipschitz. For details, see Ghirardato and Siniscalchi (2010).

Online Appendix D shows that, given an interior act h, whether a given probability $P \in ba_1(\Sigma)$ belongs to the set $C(h)$ can be directly ascertained using the DM's preferences, without invoking Theorems 6 or 7.

Finally, for preferences that satisfy the 'Weak Certainty Independence' axiom of Maccheroni et al. (2006) (e.g. multiplier, variational, or vector expected utility preferences), and under additional regularity conditions (in particular, concavity or continuous differentiability of I suffice), the sets $C(h)$ *pin down the preference* $\succcurlyeq$ *uniquely*. This follows from non-smooth analogs of the Fundamental Theorem of Calculus (cf. Ngai, Luc, and Théra, 2000). We leave a fuller investigation of this fact to future research.

A Proofs of the main results

A.1 Proof of Theorems 6 and 7, and Corollary 4

Throughout, $\succcurlyeq$ is an MBL preference with representation (I, u) and relevant priors C. For any $D \subset ba_1(\Sigma)$, we also use the notation $f \succcurlyeq_D g$ to mean that $P(u \circ f) \geq P(u \circ g)$ for all $P \in D$.

We use freely the following facts. (i) Since I is monotonic, $\partial I(u \circ h)$ consists of positive linear functionals (Rockafellar, 1980, Thm.6 Cor. 3), and consequently $a \mapsto I^\circ(c; a)$ is monotonic. (ii) $a \mapsto I^\circ(u \circ h; a)$ is continuous by Rockafellar (1980), Cor. 1 p. 268.

Lemma 9 *C is the smallest weak* compact, convex set $D \subset ba_1(\Sigma)$ such that, for all $f, g \in \mathscr{F}$, $f \succcurlyeq_D g$ implies $f \succcurlyeq g$.*

Proof: That C satisfies this property is clear, because $f \succcurlyeq_C g$ implies $f \succcurlyeq^* g$ by Proposition 1, and hence $f \succcurlyeq g$. Now suppose another set $D \subset ba_1(\Sigma)$ also satisfies this property. If $f \succcurlyeq_D g$, then, for all $\lambda \in (0,1]$ and $h \in \mathscr{F}$, also $\lambda f + (1-\lambda)h \succcurlyeq_D \lambda g + (1-\lambda)h$. Then, by assumption, $\lambda f + (1-\lambda)h \succcurlyeq \lambda g + (1-\lambda)h$ for all $\lambda \in (0,1]$ and $h \in \mathscr{F}$: that is, $f \succcurlyeq^* g$. But by Prop. 1, this implies that $f \succcurlyeq_C g$. By Prop. A.1 in GMM, this implies that $C \subset \overline{\text{co}}\, D$. ■

Lemma 10 *$f \succcurlyeq_h^* g$ implies that, for all $\mu \in (0,1]$ and $c \in B_0(\Sigma)$, $I^\circ(u \circ h; \mu u \circ f + (1-\mu)c) \geq I^\circ(u \circ h; \mu u \circ g + (1-\mu)c)$.*[11]

Proof: (Step 1) Fix $(\lambda_n), (h_n)$ as in Def. 4. Functionally, Eq. (7) is equivalent to $I(\lambda_n u \circ f + (1-\lambda_n) u \circ h_n) \geq I(\lambda_n u \circ g + (1-\lambda_n) u \circ h_n)$ eventually; in other words,

$$I(\lambda_n u \circ f + c_n) \geq I(\lambda_n u \circ g + c_n) \qquad \text{eventually}$$

for all sequences $(\lambda_n) \downarrow 0$ and (c_n) such that $c_n = (1-\lambda_n) u \circ h_n$ for some sequence $(h_n) \subset \mathscr{F}$ with $h_n \to h$.

(Step 2) For any sequence $(\lambda_n) \downarrow 0$ and $(c_n) \to u \circ h$, and for any $\mu \in (0,1]$ and $c \in B_0(\Sigma)$,

$$\lambda_n[\mu u \circ f + (1-\mu)c] + c_n = (\lambda_n \mu) u \circ f + \left[\lambda_n (1-\mu)c + c_n\right] \equiv (\lambda_n \mu) u \circ f + d_n,$$

and analogously $\lambda_n[\mu u \circ g + (1-\mu)c] + c_n = (\lambda_n \mu) u \circ g + d_n$. Since $c_n \to u \circ h$, eventually $(1-\lambda_n \mu)^{-1} d_n \in \text{int } B_0(\Sigma, u(X))$ because h is interior, $\lambda_n(1-\mu)c \to 0$, and $1 - \lambda_n \mu \to 1$; also, $d_n \to u \circ h$. Therefore, there is a sequence $(h_n) \subset \mathscr{F}$ such that $(1-\lambda_n \mu) u \circ h_n = d_n$; this sequence necessarily satisfies $h_n \to h$, and so, by Step 1, eventually

$$I(\lambda_n[\mu u \circ f + (1-\mu)c] + c_n) = I\big((\lambda_n \mu) u \circ f + d_n\big) \geq I\big((\lambda_n \mu) u \circ g + d_n\big) = I(\lambda_n[\mu u \circ g + (1-\mu)c] + c_n).$$

Subtracting $I(c_n)$ from both sides and dividing by $\lambda_n > 0$ yields

$$\frac{I(\lambda_n[\mu u \circ f + (1-\mu)c] + c_n) - I(c_n)}{\lambda_n} \geq \frac{I(\lambda_n[\mu u \circ g + (1-\mu)c] + c_n) - I(c_n)}{\lambda_n} \quad \text{eventually}$$

for all $(\lambda_n) \downarrow 0$, $\mu \in (0,1]$, $c \in B_0(\Sigma)$ and $(c_n) \to u \circ h$.

(Step 3) Finally, fix μ, c, and $\epsilon > 0$. By the definition of $I^\circ(u \circ h; \mu u \circ g + (1-\mu)c)$, there are sequences $(\lambda_n) \subset [0,1]$, $(c_n) \subset B_0(\Sigma, u(X))$ such that $\lambda_n \downarrow 0$, $c_n \to u \circ h$, and $\lim_n \frac{I(\lambda_n[\mu u \circ g + (1-\mu)c] + c_n) - I(c_n)}{\lambda_n} \geq I^\circ(u \circ h; \mu u \circ g + (1-\mu)c) - \epsilon$. Taking a subsequence if necessary,[12] it follows from Step 2 that $\lim_n \frac{I(\lambda_n[\mu u \circ f + (1-\mu)c] + c_n) - I(c_n)}{\lambda_n} \geq I^\circ(u \circ h; \mu u \circ g + (1-\mu)c) - \epsilon$. This implies that $I^\circ(u \circ h; \mu u \circ f + (1-\mu)c) \geq I^\circ(u \circ h; \mu u \circ g + (1-\mu)c) - \epsilon$. Since $\epsilon > 0$ was arbitrary, the claim follows. ■

---

[11] By Lemma 5, $\succcurlyeq_h^*$ is independent. Thus, just showing that $f \succcurlyeq_h^* g$ implies $I^\circ(u \circ h; u \circ f) \geq I^\circ(u \circ h; u \circ g)$ would be enough to establish the claim in this Lemma for $c \in B_0(\Sigma, u(X))$. However, the proof of Theorem 6 requires that the claim hold for all $c \in B_0(\Sigma)$.

[12] The sequence $\frac{I(\lambda_n[\mu u \circ f + (1-\mu)c] + c_n) - I(c_n)}{\lambda_n}$ may fail to converge. However, since $I^\circ(u \circ h; \mu u \circ f + (1-\mu)c) < \infty$ as I is locally Lipschitz, this sequence must be bounded and hence contain a convergent subsequence.

Corollary 11 *If $\partial I(u \circ h) \neq \{Q_0\}$ (the zero measure), then $\succsim_h^*$ agrees with $\succsim$ on X.*

Proof: By monotonicity of $\succsim$, $x \succsim y$ implies $x \succsim_h^* y$, so it is enough to prove that $x \succ y$ implies that $y \succsim_h^* x$ does not hold. By contradiction, suppose that $x \succ y$ (hence, $x \succsim_h^* y$) and $y \succsim_h^* x$. Then, by Lemma 10, for every $c \in B_0(\Sigma)$ and $\mu \in (0,1]$, $y \succsim_h^* x$ implies $I^\circ(u \circ h; \mu u(y) + (1-\mu)c) \geq I^\circ(u \circ h; \mu u(x) + (1-\mu)c)$. Now let $c = 1_S$ and choose $\mu > 0$ small enough so that $\alpha \equiv \mu u(x) + (1-\mu) > 0$ and $\beta \equiv \mu u(y) + (1-\mu) > 0$. Then $I^\circ(u \circ h; \alpha) = \max_{Q \in \partial I(u \circ h)} \alpha Q(S) = \alpha \max_{Q \in \partial I(u \circ h)} Q(S) \equiv \alpha M$ and similarly $I^\circ(u \circ h; \beta) = \beta M$, because $\alpha, \beta > 0$. By assumption Q_0 is not the only functional in $\partial I(u \circ h)$, and therefore, since I is monotonic, $M > 0$. But then $\alpha \leq \beta$, which contradicts the fact so $u(x) > u(y)$. ■

Lemma 12 *Assume that I is nice at $u \circ h$. For any pair $f, g \in \mathscr{F}$, $f \succsim_{C(h)} g$ implies that $f' \succsim_h^* g'$ for any spread (f', g') of (f, g).*

Proof: The claim is vacuously true if f or g are not interior acts, because in this case there is no spread of (f, g). Thus, consider a spread (f', g') of an interior pair of acts (f, g). Then there is $\epsilon > 0$ such that $u \circ f' \geq u \circ f + \epsilon$ and $u \circ g' \leq u \circ g - \epsilon$. Thus, $P(u \circ f') > P(u \circ g')$ for all $P \in C(h)$.

Suppose there are sequences $(\lambda_n) \subset [0,1]$ and $(h^n) \subset \mathscr{F}$ such that $\lambda_n \downarrow 0$, $h_n \to h$ and, by taking subsequences if necessary, $\lambda^n f' + (1-\lambda^n) h^n \prec \lambda^n g' + (1-\lambda^n) h^n$ for all n. Passing to the functional representation of $\succsim$, $I(\lambda^n u \circ f' + (1-\lambda^n) u \circ h^n) < I(\lambda^n u \circ g' + (1-\lambda^n) u \circ h^n)$ for all n.

Let $c^n = \lambda^n u \circ f' + (1-\lambda^n) h^n$, so $\lambda^n u \circ g' + (1-\lambda^n) u \circ h^n = c^n + \lambda^n [u \circ g' - u \circ f']$, and $c^n \to u \circ h$. Then, $I(c^n) < I(c^n + \lambda^n [u \circ g' - u \circ f'])$ for all n, so $\frac{I(c^n + \lambda^n [u \circ g' - u \circ f']) - I(c^n)}{\lambda^n} > 0$ for all n.

It follows that $\max_{Q \in \partial I^\circ(u \circ h)} Q(u \circ g' - u \circ f') = I^\circ(u \circ h; u \circ g' - u \circ f') \geq 0$. Hence, since I is nice at $u \circ h$, there exists $Q \neq Q_0$ in $\partial I(u \circ h)$ such that $Q(u \circ g') \geq Q(u \circ f')$; then, for $P = \frac{Q}{Q(S)} \in C(h)$, $P(u \circ g') \geq P(u \circ f')$: contradiction. ■

Proof of Theorem 6: (1): If $\partial I(u \circ h) = \{Q_0\}$, then $C(h) = \emptyset$, so the assertion holds vacuously. Thus, assume henceforth that there is $Q \neq Q_0$ such that $Q \in \partial I(u \circ h)$.

Define a relation $\geq_h^*$ on $B_0(\Sigma)$ by letting $a \geq_h^* b$ iff $I^\circ(u \circ h; \lambda a + (1-\lambda) c) \geq I^\circ(u \circ h; \lambda b + (1-\lambda) c)$ for all $\lambda \in (0,1]$ and $c \in B_0(\Sigma)$. Since the map $a \mapsto I^\circ(u \circ h; a)$ is monotonic and continuous, adapting the proof of in Prop. 4 in GMM one can easily show that $\geq_h^*$ is monotonic, reflexive, transitive, continuous (if $a^n \to a$, $b^n \to b$ and $a^n \succeq_h^* b^n$, then $a \succeq_h^* b$) and conic (if $a \succeq_h^* b$ then $\lambda a+(1-\lambda)c \succeq_h^* \lambda b+(1-\lambda)c$ for all $\lambda \in (0,1]$ and $c \in B_0(\Sigma)$). Finally, let $\alpha > \beta > 0$, and suppose that $\succeq_h^*$ is trivial. Then, since $\alpha \succeq_h^* \beta$ by monotonicity, we must have $\beta \succeq_h^* \alpha$, so for all $\lambda \in (0,1]$ and $c \in B_0(\Sigma)$, $I^\circ(u\circ h; \lambda\beta+(1-\lambda)c) \geq I^\circ(u\circ h; \lambda\alpha+(1-\lambda)c)$. Take $c = 1_S$ and any $\lambda \in (0,1]$; then $\lambda\alpha+1-\lambda, \lambda\beta+1-\lambda > 0$, so $I^\circ(u\circ h; \lambda\beta+(1-\lambda)) = [\lambda\beta+(1-\lambda)]M$ and $I^\circ(u\circ h; \lambda\alpha+(1-\lambda)) = [\lambda\alpha+(1-\lambda)]M$, where $M = \max_{Q\in\partial I(u\circ h)} Q(S) > 0$ because $\partial I(u\circ h)$ contains positive functionals other than Q_0. Hence, $\beta \succeq_h^* \alpha$ requires $\lambda\beta+(1-\lambda) \geq \lambda\alpha+(1-\lambda)$, a contradiction. Thus $\succeq_h^*$ is non-trivial.

Prop. A.2 in GMM yields a unique weak*-compact, convex set $C(h) \subset ba_1(\Sigma)$ such that $a \succeq_h^* b$ iff $P(a) \geq P(b)$ for all $P \in C(h)$. We claim that $C(h) = \overline{\text{co}}\left\{\frac{Q}{Q(S)} : Q \in \partial I(u\circ h), Q(S) > 0\right\} \equiv D(h)$.

First, we show that $C(h)$ is the smallest weak*-compact, convex set $D \subset ba_1(\Sigma)$ with the following property, henceforth (P): $P(a) \geq P(b)$ for all $P \in D$ implies $I^\circ(u\circ h; a) \geq I^\circ(u\circ h; b)$. Clearly, $C(h)$ satisfies (P), so consider another set D that also satisfies (P). If $P(a) \geq P(b)$ for all $P \in D$, then, for all $\lambda \in (0,1]$ and $c \in B_0(\Sigma)$, also $P(\lambda a+(1-\lambda)c) \geq P(\lambda b+(1-\lambda)c)$ for all $P \in D$, so by assumption $I^\circ(u\circ h; \lambda a+(1-\lambda)c) \geq I^\circ(u\circ h; \lambda b+(1-\lambda)c)$. But this means that $a \succeq_h^* b$. In other words, the relation $\succeq_D$, defined by $a \succeq_D b$ iff $P(a) \geq P(b)$ for all $P \in D$, is a subset of $\succeq_h^*$. By Prop. A.1 in GMM, $C(h) \subset \overline{\text{co}}\, D$, as claimed.

We now show that $D(h)$ is *also* the smallest weak* compact convex set that satisfies (P), which obviously implies the claim. First, suppose that, $P(a) \geq P(b)$ for all $P \in D(h)$, so $P(b-a) \leq 0$ for all $P \in D(h)$. Then also $Q(b-a) \leq 0$ for all $Q \in \partial I(u\circ h)$ [this is trivially true for $Q = Q_0$, in case $Q_0 \in \partial I(u\circ h)$]. Hence $I^\circ(u\circ h; b) = I^\circ(u\circ h; a+(b-a)) \leq I^\circ(u\circ h; a) + I^\circ(u\circ h; b-a) \leq I^\circ(u\circ h; a)$, because $I^\circ(u\circ h; b-a) = \sup_{Q\in\partial I(u\circ h)} Q(b-a) \leq 0$. Thus, $D(h)$ satisfies (P).

Let $D \subset ba_1(\Sigma)$ be another weak* compact, convex set that satisfies (P). Suppose there is $P \in D(h) \setminus D$. By the Separating Hyperplane theorem[13], there is $a \in B_0(\Sigma)$ and $\alpha \in \mathbb{R}$ such that $P(a) > \alpha$ and $P'(a) \leq \alpha$ for all $P' \in D$. Letting $b = a - \alpha$, we have $P(b) > 0$ and $P'(b) \leq 0$ for all $P' \in D$. By assumption, $P'(b) \leq 0$ for all $P' \in D$ implies $I^\circ(u\circ h; 0) \geq I^\circ(u\circ h; b)$; however,

[13] E.g. Aliprantis and Border (2007), Corollary 5.80 and Theorem 5.93. Note that, since the topologies $\sigma(ba(\Sigma), B(\Sigma))$ and $\sigma(ba(\Sigma), B_0(\Sigma))$ coincide on $ba_1(\Sigma)$ (Maccheroni et al., 2006, Appendix A), we can restrict attention to $\sigma(ba(\Sigma), B_0(\Sigma))$-continuous linear functionals on $ba(\Sigma)$.

$P(b) > 0$ implies that there is[14] $Q \in \partial I(u \circ h)$ with $Q(S) \neq 0$ and $Q(b) > 0$, so $I^\circ(u \circ h; b) = \sup_{Q' \in \partial I(u \circ h)} Q'(b) \geq Q(b) > 0 = \sup_{Q' \in \partial I(u \circ h)} 0 \cdot Q'(S) = I^\circ(u \circ h; 0)$: contradiction. Thus, $D(h) \subset D$.

To complete the proof, assume that $f \succcurlyeq_h^* g$. Then, by Lemma 10, $u \circ f \geq_h^* u \circ g$; hence $P(u \circ f) \geq P(u \circ g)$ for all $P \in C(h) = D(h)$.

(2): Suppose that $\succcurlyeq_D$ is another Bewley refinement of $\succcurlyeq_h^*$; recall that by definition $\succcurlyeq_D$ on X is represented by u, but by Corollary 11, this must be true for any Bewley refinement of $\succcurlyeq_h^*$, because I is nice at $u \circ h$ and so a fortiori $\partial I(u \circ h) \neq \{Q_0\}$. Fix an interior act $f \in \mathscr{F}$, and let x be such that $u(x) = \max_{P \in C(h)} P(u \circ f)$. We will show that $u(x) \geq \max_{P \in D} P(u \circ f)$.

Clearly, $P(u(x)) \geq P(u \circ f)$ for all $P \in C(h)$. Since f is interior, and by monotonicity so is x, there are $\epsilon > 0$, $y \in X$, $g \in \mathscr{F}$ with $u(y) = u(x) + \epsilon$ and $u \circ g = u \circ f - \epsilon$. Indeed, for all $\delta \in (0, \epsilon)$, there are y_δ and g_δ such that $u(y_\delta) = u(x) + \delta$ and $u \circ g_\delta = u \circ f - \delta$. For each such $\delta \in (0, \epsilon)$, (y_δ, g_δ) is a spread of (x, f). By Lemma 12, it must then be the case that $y_\delta \succcurlyeq_h^* g_\delta$.

Since $\succcurlyeq_D$ extends $\succcurlyeq_h^*$, conclude that $y_\delta \succcurlyeq_D g_\delta$ for all $\delta \in (0, \epsilon)$; hence, $u(y_\delta) = P(u(y_\delta)) \geq P(u \circ g_\delta)$ for all $P \in D$. Therefore, $u(x) + \delta \geq P(u \circ f) - \delta$ for all $P \in D$ and all $\delta \in (0, \epsilon)$. It follows that $u(x) \geq P(u \circ f)$ for all $P \in D$, i.e. $u(x) \geq \max_{P \in D} P(u \circ f)$, as claimed.

To sum up, for all interior $f \in \mathscr{F}$, $\max_{P \in C(h)} P(u \circ f) \geq \max_{P \in D} P(u \circ f)$. Since any $a \in B_0(\Sigma)$ can be written as $\alpha u \circ f + \beta$ for some $f \in \mathscr{F}^{\text{int}}$ and $\alpha, \beta \in \mathbb{R}$, $\max_{P \in C(h)} P(a) \geq \max_{P \in D} P(a)$ for all $a \in B_0(\Sigma)$. By standard results (e.g. Aliprantis and Border, 2007, Theorem 7.51), this implies that $D \subset C(h)$, i.e. $\succcurlyeq_D$ is a richer Bewley relation than $\succcurlyeq_h^*$. ■

Proof of Theorem 7: fix an interior pair (f, g). Assume that (1) holds, and fix $\epsilon > 0$ such that $u \circ f + \epsilon, u \circ g - \epsilon \in B_0(\Sigma, u(X))$. Then, for all $\delta \in (0, \epsilon)$, there exist $f_\delta, g_\delta \in \mathscr{F}$ with $u \circ f_\delta = u \circ f + \delta$ and $u \circ g_\delta = u \circ g - \delta$; note that (f_δ, g_δ) is a spread of (f, g). Then $f_\delta \succcurlyeq_h^* g_\delta$, so by (1) in Theorem 6, $P(u \circ f) + \delta = P(u \circ f_\delta) \geq P(u \circ g_\delta) = P(u \circ g) - \delta$ for all $P \in C(h)$ and all $\delta \in (0, \epsilon)$. Therefore, (2) with $D = C(h)$ follows.

The converse, again with $D = C(h)$, is established in Lemma 12. Finally, suppose there is another set D for which (1) and (2) are equivalent (again using utility u in view of Corollary 11).

[14]This is immediate if $P = \frac{Q}{Q(S)}$ for some $Q \in \partial I(u \circ h)$. If not, there is a net (P_ι) in $\mathrm{co}\{\frac{Q}{Q(S)} : Q \in \partial I(u \circ h), Q(S) > 0\}$ that converges to P in the weak* topology. Then, $P_\iota(b) \to P(b)$, so there is $\bar{\iota}$ such that $P_\iota(b) > 0$ for all ι following $\bar{\iota}$. Since any such P_ι is a convex combination of elements of $\{\frac{Q}{Q(S)} : Q \in \partial I(u \circ h), Q(S) > 0\}$, the claim follows.

Consider a pair (f,g) of interior acts. Suppose that $f \succcurlyeq_{C(h)} g$: then (2) holds for set $C(h)$, hence (1) must hold. But by assumption this implies that (2) must hold for set D as well, and therefore $f \succcurlyeq_D g$. Since Bewley prefences satisfy Independence, $\succcurlyeq_{C(h)} \subset \succcurlyeq_D$. By the same argument, $\succcurlyeq_D \subset \succcurlyeq_{C(h)}$. It follows that $D = C(h)$. ■

A.2 Proof of Theorem 2 and Corollary 8

We must show that C is the closed convex hull of all $C(h)$, for $h \in \mathscr{F}^{\text{int}}$.

Claim: for all $f,g \in \mathscr{F}$, $f \succcurlyeq_{C(h)} g$ for all $h \in \mathscr{F}^{\text{int}}$ implies $f \succcurlyeq g$.

Proof: assume first that f and g are interior. By Lebourg's Mean Value Theorem (Lebourg, 1979, Theorem 1.7), there is $\mu \in (0,1)$ and $Q \in \partial I(\mu u \circ f + (1-\mu) u \circ g)$ such that $I(u \circ f) - I(u \circ g) = Q(u \circ f) - Q(u \circ g)$. Since $\mu f + (1-\mu)g$ is interior, the assumption that $P(u \circ f) \geq P(u \circ g)$ for all $P \in C(\mu f + (1-\mu)g)$ implies that $Q(u \circ f) \geq Q(u \circ g)$ [if $Q = Q_0$ this is trivially true]. Hence, $I(u \circ f) \geq I(u \circ g)$, i.e. $f \succcurlyeq g$, as claimed. If now f,g are not interior, pick x interior and consider $\lambda x + (1-\lambda) f, \lambda x + (1-\lambda) g$. If $P(u \circ f) \geq P(u \circ g)$ for all interior h and all $P \in C(h)$, then also $P(\lambda u(x) + (1-\lambda) u \circ f) \geq P(\lambda u(x) + (1-\lambda) u \circ g)$ for all such h, P. As was just shown, this implies $\lambda x + (1-\lambda) f \succcurlyeq \lambda x + (1-\lambda) g$. Since this holds for all λ, continuity yields $f \succcurlyeq g$, as required.

By Lemma 9, this Claim implies that $C \subset \overline{\text{co}} \bigcup_{h \in \mathscr{F}^{\text{int}}} C(h)$. Conversely, suppose $f \succcurlyeq_C g$. Then $f \succcurlyeq^* g$; in particular, for every $h \in \mathscr{F}^{\text{int}}$, $f \succcurlyeq^*_h g$. But then, Part (1) of Theorem 6 shows that $f \succcurlyeq_{C(h)} g$. Applying Prop. A.1 in GMM to the Bewley preference $\succcurlyeq_{C(h)}$ now implies that $C(h) \subset C$.

Note that the above also shows: $f \succcurlyeq_{C(h)} g$ for all interior h if and only if $f \succcurlyeq^* g$. Since $f \succcurlyeq^* g$ directly and trivially implies that $f \succcurlyeq^*_h g$, and Part 1 of Theorem 6 shows that $f \succcurlyeq^*_h g$ implies $f \succcurlyeq_{C(h)} g$, we can also conclude that $f \succcurlyeq^* g$ if and only if $f \succcurlyeq^*_h g$ for all interior h.

References

Charalambos Aliprantis and Kim Border. *Infinite Dimensional Analysis.* Springer Verlag, Berlin, 2007.

Aurélien Baillon, Olivier L'Haridon, and Laetitia Placido. Ambiguity models and the machina paradoxes. *American Economic Review*, forthcoming.

Truman Bewley. Knightian decision theory: Part I. *Decisions in Economics and Finance*, 25(2): 79–110, November 2002. (first version 1986).

S. Cerreia-Vioglio, P. Ghirardato, F. Maccheroni, M. Marinacci, and M. Siniscalchi. Rational preferences under ambiguity. *Economic Theory*, pages 1–35, 2011.

Simone Cerreia-Vioglio, Fabio Maccheroni, Massimo Marinacci, and Luigi Montrucchio. Uncertainty averse preferences. Carlo Alberto Notebook 77, 2008.

Alain Chateauneuf and Josè Heleno Faro. Ambiguity through confidence functions. *Journal of Mathematical Economics*, 45(9-10):535–558, September 2009.

Frank H. Clarke. *Optimization and Nonsmooth Analysis*. J. Wiley, New York, 1983.

J. Dow and S.R. da Costa Werlang. Uncertainty aversion, risk aversion, and the optimal choice of portfolio. *Econometrica: Journal of the Econometric Society*, pages 197–204, 1992.

Larry G. Epstein. A definition of uncertainty aversion. *Review of Economic Studies*, 66:579–608, 1999.

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

Paolo Ghirardato and Marciano Siniscalchi. A more robust definition of multiple priors. Mimeo, June 2010.

Paolo Ghirardato, Fabio Maccheroni, and Massimo Marinacci. Differentiating ambiguity and ambiguity attitude. *Journal of Economic Theory*, 118:133–173, 2004.

Itzhak Gilboa and David Schmeidler. Maxmin expected utility with a non-unique prior. *Journal of Mathematical Economics*, 18:141–153, 1989.

Itzhak Gilboa, Fabio Maccheroni, Massimo Marinacci, and David Schmeidler. Objective and subjective rationality in a multiple prior model. *Econometrica*, 78:755–770, 2010.

Simon Grant and Ben Polak. Mean-dispersion preferences and absolute uncertainty aversion. Technical Report 1805, Cowles Foundation Discussion Paper, 2011a.

Simon Grant and Ben Polak. A two-parameter model of dispersion aversion. 2011b.

Lars P. Hansen and Thomas J. Sargent. Robust control and model uncertainty. *The American Economic Review*, 91:60–66, 2001.

Chip Heath and Amos Tversky. Preference and belief: Ambiguity and competence in choice under uncertainty. *Journal of Risk and Uncertainty*, 4:5–28, 1991.

Peter Klibanoff, Massimo Marinacci, and Sujoy Mukerji. A smooth model of decision making under ambiguity. *Econometrica*, 73:1849–1892, 2005.

Peter Klibanoff, Sujoy Mukerji, and Kyoungwon Seo. Relevance and symmetry. Discussion Paper n. 539, Oxford University, June 2011.

G. Lebourg. Generic differentiability of Lipschitzian functions. *Transactions of the American Mathematical Society*, pages 125–144, 1979.

Fabio Maccheroni, Massimo Marinacci, and Aldo Rustichini. Ambiguity aversion, robustness, and the variational representation of preferences. *Econometrica*, 74:1447–1498, 2006.

M.J. Machina. 'Expected utility' analysis without the independence axiom. *Econometrica*, pages 277–323, 1982.

M.J. Machina. 'Expected utility/subjective probability' analysis without the sure-thing principle or probabilistic sophistication. *Economic Theory*, 26(1):1–62, 2005.

Klaus Nehring. Imprecise probabilistic beliefs as a context for decision-making under ambiguity. Mimeo, November 2002.

HV Ngai, DT Luc, and M. Théra. Approximate convex functions. *J. Nonlinear Convex Anal*, 1(2): 155–176, 2000.

L. Rigotti, C. Shannon, and T. Strzalecki. Subjective beliefs and ex ante trade. *Econometrica*, 76 (5):1167–1190, 2008.

A.W. Roberts and D.E. Varberg. Another proof that convex functions are locally lipschitz. *The American Mathematical Monthly*, 81(9):1014–1016, 1974.

R. Tyrrell Rockafellar. Generalized directional derivatives and subgradients of nonconvex functions. *Canadian Journal of Mathematics*, XXXII:257–280, 1980.

David Schmeidler. Subjective probability and expected utility without additivity. *Econometrica*, 57:571–587, 1989.

Marciano Siniscalchi. A behavioral characterization of plausible priors. *Journal of Economic Theory*, 128:91–135, 2006.

Marciano Siniscalchi. Vector expected utility and attitudes toward variation. *Econometrica*, 77: 801–855, 2009.