# Economics Department Discussion Papers Series

ISSN 1473 – 3307

## THE $\alpha$-MEU MODEL: A COMMENT

Jürgen Eichberger, Simon Grant, David Kelsey and Gleb A. Koshevoy

**Paper number 09/03**

URL: http://business-school.exeter.ac.uk/economics/papers/

# THE $\alpha$-MEU MODEL: A COMMENT[1]


Jürgen Eichberger
Alfred Weber Institut,
Universität Heidelberg, Germany.

Simon Grant
Rice University,
Texas, USA

David Kelsey
Department of Economics,
University of Exeter, England.

Gleb A. Koshevoy
Central Institute of Mathematics and Economics RAS,
Nahimovskii pr.47,
117418 Moscow, Russia


11th September 2009


[1]Research in part supported by ESRC grant no. RES-000-22-0650. For comments and discussion we would like to thank Dieter Balkenborg, Paolo Ghirardato, Klaus Nehring, Jan Wenzelburger, Yiannis Vailakis, participants in seminars in Exeter and Heidelberg the referee and associate editor of this journal.

**Abstract**

In [7] Ghirardato, Macheroni and Marinacci (GMM) propose a method for distinguishing between perceived ambiguity and the decision-maker's reaction to it. They study a general class of preferences which they refer to as invariant biseparable. This class includes CEU and MEU. They axiomatize a subclass of $\alpha$-MEU preferences. If attention is restricted to finite state spaces, we show that any $\alpha$-MEU preference relation, satisfies GMM's axioms *if and only if* $\alpha = 0$ or 1, that is, the preferences must be either maxmin or maxmax. We show by example that these axioms may be satisfied when the state space is [0,1].

**Keywords**, Ambiguity, multiple priors, invariant biseparable, Clarke derivative, ambiguity-preference.

**JEL Classification:** D81.

**Address for Correspondence** Simon Grant, Department of Economics, Rice University, PO Box 1892, Houston, Texas 77251-1892, USA (e-mail: sgrant@rice.edu)

# 1 Introduction

Ghirardato, Macheroni and Marinacci [7] (henceforth GMM), axiomatize a class of preferences they refer to as *invariant biseparable*, that encompasses both the Choquet expected utility (henceforth CEU) model of Schmeidler [12] and the maxmin expected utility (also known as the multiple prior) model of Gilboa & Schmeidler [8].[1] They define for a given preference relation $\succcurlyeq$ from this class, the (generally) partial ordering $\succcurlyeq^*$ that is the maximal sub-relation of $\succcurlyeq$ satisfying all the axioms of subjective expected utility (SEU) except completeness.[2] They refer to $\succcurlyeq^*$ as the *unambiguous preference relation* and show that it admits a representation in the style of Bewley [2]: in particular, there is a utility function $u(\cdot)$ defined on the set of outcomes $X$ and a non-empty, compact and convex set of probability measures $\mathcal{D}$ defined on the state space $S$ such that for any pair of acts $f$ and $g$,

$$f \succcurlyeq^* g \iff \int_S u(f(s))\, dP(s) \geqslant \int_S u(g(s))\, dP(s), \forall\ P \in \mathcal{D}.$$

The relation $\succcurlyeq^*$ is complete if and only if $\mathcal{D}$ is a singleton in which case $\succcurlyeq$ equals $\succcurlyeq^*$ and has the SEU form.

Furthermore, they establish the existence of a function $\beta(\cdot)$ that maps each act $f$ to a weight $\beta(f)$ in [0, 1], such that $\succcurlyeq$ can be represented by the functional:

$$V(f) = \beta(f) \min_{P\in\mathcal{D}} \int_S u(f(s))\, dP(s) + (1-\beta(f)) \max_{P\in\mathcal{D}} \int_S u(f(s))\, dP(s). \tag{1}$$

Typically $V$ will have kinks at constant acts, (that is, acts which assign the same utility to every state). Thus it is not possible to apply conventional notions of differentiation. Instead GMM use the Clarke derivative.[3] They also show that the set $\mathcal{D}$ admits a straightforward differential characterization. For our purposes it is enough to note that for the functional $I : \mathbb{R}^S \to \mathbb{R}$, in which $I(u \circ f) = V(f)$, for each act $f$, the set $\mathcal{D}$ is precisely $\partial I(0)$, the Clarke differential of $I$ at 0.

GMM are careful to note the following feature of their representation. To generate preferences in their class it is not enough to fix an arbitrary (non-empty, weak* compact and convex) set of probability

[1] See GMM [7], Axioms 1-5. The formal statements of these axioms appear in the appendix below.

[2] Note this is equivalent but not identical to the original definition, for details see [7]. Related research can be found in Nehring [11].

[3] The Clarke derivative is one way to extend the concept of a derivative on $\mathbb{R}^n$ to functions which have some kinks. It can also be seen as a generalization of the super-gradient to a function which is not necessarily concave. For further details see Clarke [4].

measures $\mathcal{D}$ and an arbitrary index $\beta(\cdot)$ and substitute them into equation (1). Rather in order for expression (1) to generate an invariant biseparable preference relation for a given set $\mathcal{D}$ and index $\beta(.)$, we need to check that the associated Clarke differential of $I$ at 0 is indeed equal to $\mathcal{D}$.

At first glance expression (1) appears closely related to the classic $\alpha$-MEU model:

$$V(f) = \alpha \min_{P \in \mathcal{D}} \int_S u(f(s))\, dP(s) + (1-\alpha) \max_{P \in \mathcal{D}} \int_S u(f(s))\, dP(s). \tag{2}$$

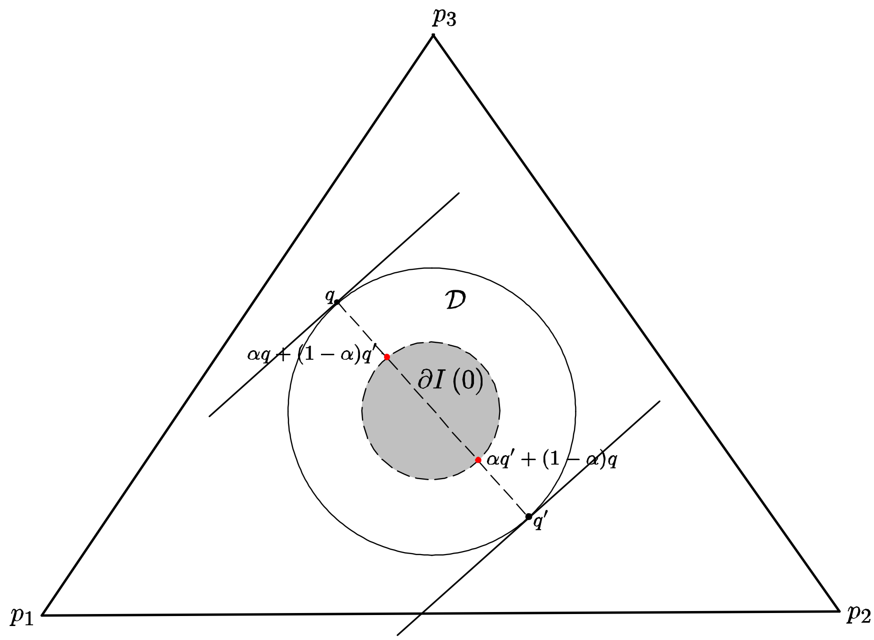
However there are two differences. First in the classic $\alpha$-MEU model the weight $\alpha$ on the minimum expected utility is constant, whereas in expression (1) the weight $\beta(f)$ depends on the act $f$. Second in the classic $\alpha$-MEU model the set $\mathcal{D}$ can be any non-empty weak* compact set of probability measures, whereas in expression (1), $\mathcal{D}$ must be equal to the Clarke differential at 0.

GMM provide an axiomatic characterization that combines the key features of the two models: the ambiguity aversion index $\beta(\cdot)$ is constant and equal to some fixed weight $\alpha$ in [0, 1], and the set of probabilities is given by the Clarke differential at 0.[4] That is, the preferences admit a representation of the form given in expression (2) with the restriction that $\mathcal{D} = \partial I(0)$, where $I(u \circ f) = V(f)$. Imposing the additional restriction that $\alpha(f)$ be constant implies GMM's representation must satisfy a type of fixed point property. If one starts with a given set $\mathcal{D}$ and constructs a set of $\alpha$-MEU preferences with this set of priors, then it is necessary that the Clarke differential at 0 be equal to $\mathcal{D}$. For a finite state space, however, we show that any relation that satisfies GMM's axiomatization (that is, their Axioms 1-5 and 7) the constant ambiguity aversion index $\alpha$ is equal either to 0 or to 1. That is, the preference relation is either maxmax expected utility or maxmin expected utility.

Our proof strategy is to fix a closed convex set, $\mathcal{D}$, of probability distributions on $S$ and an $\alpha$ in [0, 1], and consider the preferences defined by expression (2) and define $I : \mathbb{R}^n \to \mathbb{R}$ by $I(u \circ f) = V(f)$. If the preferences satisfy the GMM axioms, then $\partial I(0)$ should yield the original set $\mathcal{D}$. The analysis in section 3 shows that when we take the Clarke derivative, we do not get back the original set $\mathcal{D}$ unless the ambiguity index, $\alpha$, is equal either to 1 or to 0.

The intuition is most clear in the case where $\mathcal{D}$ is a circle as shown in figure 1. The figure considers a given act $f$. The expected value of $f$ is maximised at $q \in \mathcal{D}$ and minimised at $q'$. The decision weight on $f$ is accordingly $\alpha q' + (1-\alpha) q$. As $f$ varies the decision weights trace out the boundary of the

[4] The axiomatization consists of their Axioms 1-5 and an additional axiom Axiom 7. The formal statement of this axiom appears in the appendix below.




inner circle. The Clarke differential, $\partial I\,(0)$, is the convex hull of these points, which corresponds to the shaded area in the diagram. As can be seen, it is a proper subset of the set of priors $\mathcal{D}$.

A similar result does not hold for infinite state spaces. We show that there exist examples of $\alpha$-MEU preferences satisfying GMM's axioms.

**Organization of the paper** The next section provides a review of some of the mathematical techniques we shall be using. In section 3 we show that when the state space is finite there is no $\alpha$-MEU preference, which satisfies the GMM axioms. However there are examples of such preferences over infinite state spaces as we shall demonstrate in section 4. Formal statements of GMM's axioms appear in the appendix.

## 2 Mathematical Preliminaries

This section reviews some mathematical concepts which we need, in particular the Clarke derivative.

### 2.1 Lipschitz functions

The Clarke derivative is defined for functions which are locally Lipschitz. These are defined as follows.

**Definition 1** *Let $X$ be a subset of a Banach space. A function $f : X \to \mathbb{R}$ is said to be* Lipschitz *if there* *exists* $L > 0$ *such that for all* $x, y \in X$, $|f(x) - f(y)| < L \|x - y\|$. *A function* $g : X \to \mathbb{R}$, *is said to be* locally Lipschitz *if for all* $x \in X$, *there is a neighbourhood of* $x$ *on which g is Lipschitz.*

**Lemma 1** *Let* $f$ *and* $g$ *be two real valued functions defined on an open subset* $U$ *of* $\mathbb{R}^n$. *Then if both* $f$ *and* $g$ *are Lipschitz so is* $f - g$.

**Proof.** Since $f$ and $g$ are Lipschitz, there exist $L', L'' > 0$ such that $|f(x) - f(y)| < L' \|x - y\|$ and $|g(x) - g(y)| < L'' \|x - y\|$. Now let $L = \max\{L', L''\}$ and note that $(f - g)(x) = f(x) - g(x)$, then we have $|(f - g)(x) - (f - g)(y)| = |f(x) - g(x) - f(y) + g(y)| \leqslant |f(x) - f(y)| + |g(x) - g(y)| < L \|x - y\| + L \|x - y\| < (2L) \|x - y\|$. ■

Clarke [4] shows that any bounded convex function is Lipschitz.

**Proposition 1** *(Clarke [4], Proposition 2.2.6, p.34)* *Let* $U$ *be an open subset of a Banach space* $X$, *and let* $f : U \to \mathbb{R}$ *be convex and bounded above on a neighbourhood of some point of* $U$. *Then for any* $x$ *in* $U$, $f$ *is Lipschitz near* $x$.

### 2.2 Derivatives

The usual derivative on $\mathbb{R}^n$ is defined as follows.

**Definition 2** *A function* $V : \mathbb{R}^n \to \mathbb{R}$, *is said to be* differentiable *at* $x$ *if there exists a linear function* $dV_x : \mathbb{R}^n \to \mathbb{R}$ *such that:*

$$\lim_{h \to 0} \frac{V(x+h) - V(x) - dV_x(h)}{\|h\|} = 0.$$

The limit is required to be independent of the direction from which $h$ approaches 0. The linear function $dV_x$ may be represented by the gradient, $\nabla V$, of $V$ in the sense that $dV_x(h) = \nabla V \cdot h$, for all $h \in \mathbb{R}^n$.

Typically when there is ambiguity, preferences are represented by functions which are not differentiable everywhere. To overcome this problem GMM use the Clarke derivative. Below we define the Clarke (directional) derivative which measures the slope of a function in a particular direction.

**Definition 3** *Let* $V : \mathbb{R}^n \to \mathbb{R}$ *be a locally Lipschitz function. The* Clarke (lower) directional derivative *of* $V$ *at* $x$ *in direction* $d$ *is defined by:*

$$DV(x, d) = \liminf_{y \to x, t \downarrow 0} \frac{V(y + td) - V(y)}{t}.$$

At a point where $V$ is differentiable $DV(x,d)$ is equal to the derivative $dV_x(d)$. If $V$ is not differentiable at $x$, there is locally more than one normal vector to the indifference curves of $V$. The Clarke differential is essentially the closure of the convex hull of these local normal vectors. It can be seen as playing the role of the normal vector at points where the function is not differentiable.

**Definition 4** *Let $V : \mathbb{R}^n \to \mathbb{R}$ be a locally Lipschitz function. The* Clarke differential *of $V$ at $x$ is defined by:*

$$\partial V(x) = \left\{z \in \mathbb{R}^n : z.d \geqslant DV(x,d), \forall d \in \mathbb{R}^n\right\}.$$

The Clarke differential is a generalization of the derivative on $\mathbb{R}^n$. Recall that at a point where a function is differentiable, the derivative may be represented by the gradient vector. The Clarke differential is equal to the gradient at points where the function is differentiable. A Lipschitz function on $\mathbb{R}^n$ is differentiable almost everywhere. Let $\hat{y}$ be a point where $V$ is not differentiable. Then there exists a sequence of points, at which $V$ is differentiable, which tends to $\hat{y}$. One can then consider the limit of the gradient of $V$ at these points. In general, the limit will depend on the sequence chosen. Thus we get a *set* of gradients at $\hat{y}$, which is the union of the limits of the gradients taken over all sequences which converge to $\hat{y}$. The Clarke differential is the convex hull of this set of gradients. The following result characterizes the Clarke differential in finite dimensional spaces. Its proof can be found in Clarke [4].

**Theorem 1** *(Clarke [4], Theorem 2.5.1) Let $V : \mathbb{R}^n \to \mathbb{R}$ be Lipschitz near $x$ and suppose $N$ is any null set (i.e. a set of Lebesgue measure 0) in $\mathbb{R}^n$. Then*

$$\partial V(x) = \operatorname{co}\left\{\lim \nabla V(x_i) : x_i \to x, x_i \in \Gamma_V, x_i \notin N\right\},$$

*where $\Gamma_V$ denotes the set of points at which $V$ is differentiable and $\operatorname{co}(A)$ denotes the convex hull of $A$.*

## 3 Finite State Spaces

The main result of this section is to show that when the state space is finite, GMM's axioms 1-5 plus 7 imply that the weight $\alpha$ in expression (2) is equal either to 1 or to 0. First we shall present the proof, then we shall discuss some examples which illustrate key points.

### 3.1 The Main Result

Throughout this section we assume that there is a finite set, $S$, of $n$ states of nature. Let $\Delta(S)$ denote the set of probability distributions over $S$. For simplicity we shall also assume that acts pay-off in utility terms, hence an *act* is a function from $S$ to $\mathbb{R}$. This is without any essential loss of generality, since our analysis could also be conducted using a conventional utility function over outcomes, if desired. As a result we may identify the functional $I$ with the functional $V$ in expression (2). This allows us to write the Clarke differential at 0 as $\partial V(0)$. The set of all acts is denoted by $A(S)$, which can be identified with $\mathbb{R}^n$.

The strategy of proof is as follows. As already noted, invariant biseparable preferences represented by expression (1) must satisfy a fixed point property. Imposing the extra restriction that $\beta(f)$ be a constant function implies that this fixed point property cannot be satisfied unless $\beta(f) \equiv 1$ or $\beta(f) \equiv 0$. In particular, if $\beta(f) = \alpha$ for some $\alpha$ in $(0, 1)$, then the extreme points of the set of priors, $\mathcal{D}$, are not included in the Clarke differential.

Let $\mathcal{D}$ be a given closed convex set of probabilities on $S$ and define the functions $\phi, \psi : A(S) \to \mathbb{R}$ by $\phi(f) = \min_{p\in\mathcal{D}} \mathbf{E}_p f$ and $\psi(f) = \max_{p\in\mathcal{D}} \mathbf{E}_p f$. That is, $\phi$ and $\psi$ represent maxmin and maxmax expected utility preferences respectively. The functions $\phi$ and $\psi$ are clearly not differentiable at constant acts. If $\mathcal{D}$ does not have full dimension (that is, $n-1$) or there are kinks in the boundary of $\mathcal{D}$, they may have other points of non-differentiability as well. However, since $\phi$ is concave and $\psi$ is convex, these functions are differentiable almost everywhere.

In order to apply the analysis from [4] we need to establish that $V$ is Lipschitz, which is shown in the next result.

**Lemma 2** *For all $f \in A(S)$, $V$ is locally Lipschitz at $f$.*

**Proof.** Let $B$ denote the closed ball radius $\epsilon$ around $f$ and let $\bar{x} = \max_{s\in S} f(s)$. Then for all $g \in B$, $\phi(g) \leqslant \bar{x} + \epsilon$ and $\psi(g) \leqslant \bar{x} + \epsilon$. Hence both $\psi(f)$ and $\phi(f)$ are bounded on a neighbourhood of $f$. Both $(1-a)\psi(f)$ and $-\alpha\phi(f)$ are convex functions and are therefore locally Lipschitz by Proposition 1. Since $V$ is the difference of these two functions, which are locally Lipschitz, $V$ itself is locally Lipschitz by Lemma 1. ■

The next result shows that at a point where $\phi$ is differentiable, the derivative must be equal to the minimising probability distribution. It also finds an expression for the derivative of $V$ at points where both $\phi$ and $\psi$ are differentiable. If $f \in A(S)$ is a given act, we shall use the notation $\underline{p}^f$ (resp. $\bar{p}^f$) to denote an element of $\operatorname{argmin}_{p \in \mathcal{D}} p \cdot f$ (resp. $\operatorname{argmax}_{p \in \mathcal{D}} p \cdot f$).

**Lemma 3**

1. *Suppose that $\phi$ (resp. $\psi$) is differentiable at $f$ then $d\phi_f(y) = \underline{p}^f \cdot y$, (respectively, $d\psi_f(y) = \bar{p}^f \cdot y$) or equivalently $\underline{p}^f = \nabla \phi(f)$, (respectively, $\bar{p}^f = \nabla \psi(f)$).*

2. *If $\phi$ (resp. $\psi$) is differentiable at then* $\operatorname{argmin}_{p \in \mathcal{D}} p.f$ *(resp.* $\operatorname{argmax}_{p \in \mathcal{D}} p.f$*) is unique.*

3. *Let $V$ be an $\alpha$-MEU preference functional and suppose that $\phi$ and $\psi$ are differentiable at $f$, then $V$ is differentiable at $f \in A(S)$ and $\nabla V(f) = \alpha \underline{p}^f + (1-\alpha) \bar{p}^f$.*

**Proof.** We shall prove part 1 by contradiction. If the result is false then there exists an act $f$ such that $\phi$ is differentiable at $f$ and $d\phi_f(y) = z.y$, where $z \notin \operatorname{argmin}_{p \in \mathcal{D}} p.f$. If $h = \epsilon f$, where $\epsilon \in \mathbb{R}$, $\epsilon > 0$, then $\phi(f+h) = \underline{p}^f.(f+h)$. Thus $\frac{\phi(f+h)-\phi(f)-z.h}{\|h\|} = \frac{\underline{p}^f.(f+\epsilon f)-\underline{p}^f.f-\epsilon z.f}{\epsilon\|f\|} = \frac{\underline{p}^f.f-z.f}{\|f\|} < 0$, since $z \notin \operatorname{argmin}_{p \in \mathcal{D}} p.f$. Hence $\lim_{h \to 0} \frac{\phi(f+h)-\phi(f)-z.h}{\|h\|} \neq 0$, which contradicts the assumption that $d\phi_f(y) = z.y$. This establishes that $d\phi_f(y) = \underline{p}^f.y$. By similar reasoning we may show that $d\psi_f(y) = \bar{p}^f.y$.

To prove part 2, assume $\phi$ is differentiable at $f$ then $d\phi_f = \hat{q} \in \operatorname{argmin}_{p \in \mathcal{D}} p.f$. Suppose if possible there exists $\tilde{q} \neq \hat{q}$ such that $\tilde{q} \in \operatorname{argmin}_{p \in \mathcal{D}} p.f$. Consider $h \in \mathbb{R}^n$ such that $\tilde{q}.h = 0$, and $\hat{q}.h > 0$ and let $\bar{q} \in \operatorname{argmin}_{p \in \mathcal{D}} p.(f+\epsilon h)$. Then $\bar{q}(f+\epsilon h) \leqslant \tilde{q}(f+\epsilon h) = \hat{q}.f$. Thus $\frac{\phi(f+\epsilon h)-\phi(f)-d\phi_f(\epsilon h)}{\|\epsilon h\|} = \frac{\bar{q}(f+\epsilon h).-\hat{q}.f-\epsilon h.\hat{q}}{\epsilon\|h\|} \leqslant \frac{\tilde{q}.f+\epsilon\tilde{q}.h-\hat{q}.f}{\epsilon\|h\|} - \frac{h.\hat{q}}{\|h\|} = -\frac{h.\hat{q}}{\|h\|} < 0$. Hence $\lim_{\epsilon \to 0} \frac{\phi(f+\epsilon h)-\phi(f)-d\phi_f(\epsilon h)}{\|\epsilon h\|} < 0$. However this contradicts the assumption that $\phi$ is differentiable at $f$.

Part 3 follows from part 1 and linearity of the derivative on $\mathbb{R}^n$. ■

Let $L$ denote the linear span of $\{p - q : p, q \in \mathcal{D}\}$[5] and denote by $L^{\perp}$ the orthogonal complement of the vector space $L$. If $\mathcal{D}$ has full rank then $L^{\perp}$ will consist just of the constant vectors in $A(S) = \mathbb{R}^n$. If the dimension of $\mathcal{D}$ is less than $n-1$, then $L$ will contain, in addition, non-constant acts with respect to which $\operatorname{argmin}_{p \in \mathcal{D}} p \cdot f = \operatorname{argmax}_{p \in \mathcal{D}} p \cdot f = \mathcal{D}$ holds. Recall that any $f \in A(S)$ can be uniquely written in the form $f = g + h$, where $g \in L$ and $h \in L^{\perp}$. The next result relates the Clarke differential $\partial V(0)$ to $L$.

[5] We need to consider differences, since $\Delta(S)$ is an affine subspace not a linear subspace of $\mathbb{R}^n$.

**Lemma 4** *Let $\mathcal{D} \subseteq \Delta(S)$ be a closed convex subset with cardinality greater than* 1*; let $V : \mathbb{R}^n \to \mathbb{R}$ be an $\alpha$-MEU preference functional with set of priors $\mathcal{D}$ then,*

$$\partial V(0) \subseteq \text{co}\left\{\alpha \, \text{argmin}_{p\in\mathcal{D}} \, p \cdot f + (1-\alpha)\, \text{argmax}_{p\in\mathcal{D}} \, p \cdot f : f \in L \setminus \{0\}\right\}.$$

**Proof.** Since $\phi$ is a concave function and $\psi$ is a convex function, the set of points at which they are both differentiable, $\Gamma_\phi \cap \Gamma_\psi$, is of full Lebesgue measure. Hence by Theorem 1 and Lemma 3,

$$\partial V(0) = \text{co}\left\{\lim \nabla V(f_n) : f_n \to 0, \, f_n \in \Gamma_V \cap \Gamma_\phi \cap \Gamma_\psi\right\}$$

$= \text{co}\left\{\lim\left(\alpha \underline{p}^{f_n} + (1-\alpha)\,\bar{p}^{f_n}\right) : f_n \to 0, \, f_n \in \Gamma_V \cap \Gamma_\phi \cap \Gamma_\psi\right\}$. If $h \in L^\perp$ then $\phi$ and $\psi$ are not differentiable at $h$, because for each $h \in L^\perp$, $p \cdot h = p' \cdot h$ for all $p, p' \in \mathcal{D}$. Hence $\text{argmin}_{p\in\mathcal{D}} \, p \cdot h$ and $\text{argmax}_{p\in\mathcal{D}} \, p \cdot h$ are not singleton since $\mathcal{D}$ is not singleton.

Consider a particular sequence $f_n \to 0$, $f_n \in \Gamma_V \cap \Gamma_\phi \cap \Gamma$ such that $\lim\left(\alpha \underline{p}^{f_n} + (1-\alpha)\,\bar{p}^{f_n}\right)$ exists. Fix an $n$. Write $f_n = g_n + h_n$, where $g_n \in L$ and $h_n \in L^\perp$. Define $\hat{g}_n = \frac{g_n}{\|g_n\|}$. Since $f_n \in \Gamma_\phi \cap \Gamma_\psi$ we know that $f_n \notin L^\perp$ and hence $g_n \neq 0$. Since $p \cdot h_n = p' \cdot h_n$ for all $p, p' \in \mathcal{D}$, we have $\underline{p}^{f_n} = \underline{p}^{\hat{g}_n}$ and $\bar{p}^{f_n} = \bar{p}^{\hat{g}_n}$. Therefore $\alpha \underline{p}^{f_n} + (1-\alpha)\,\bar{p}^{f_n} = \alpha \underline{p}^{\hat{g}_n} + (1-\alpha)\,\bar{p}^{\hat{g}_n}$.

Returning to the sequence $f_n \to 0$, $f_n \in \Gamma_V \cap \Gamma_\phi \cap \Gamma_\psi$, consider the corresponding sequences of $\hat{g}_n$'s, $\underline{p}^{\hat{g}_n}$'s and $\bar{p}^{\hat{g}_n}$'s. By construction the $\hat{g}_n$'s lie in a compact set (the unit ball). The $\underline{p}^{\hat{g}_n}$'s and $\bar{p}^{\hat{g}_n}$'s also lie in a compact set (the simplex). Hence, by taking a subsequence if necessary, we may assume that the three sequences $\hat{g}_n$, $\underline{p}^{\hat{g}_n}$ and $\bar{p}^{\hat{g}_n}$ all converge. Let $\hat{g}$, $\underline{p}$ and $\bar{p}$ be the respective limit points. By construction $\hat{g} \neq 0$. Furthermore $\hat{g} \in L$ because $L$ is a finite dimensional subspace and therefore closed.

By the upper hemi-continuity of argmax and argmin we know that $\bar{p} \in \text{argmax}_{p\in\mathcal{D}} \, p \cdot \hat{g}$ and $\underline{p} \in \text{argmin}_{p\in\mathcal{D}} \, p \cdot \hat{g}$. Putting the steps together we have $\lim\left(\alpha \underline{p}^{f_n} + (1-\alpha)\,\bar{p}^{f_n}\right) = \alpha \underline{p} + (1-\alpha)\,\bar{p} \in \alpha \, \text{argmin}_{p\in\mathcal{D}} \, p \cdot \hat{g} + (1-\alpha)\, \text{argmax}_{p\in\mathcal{D}} \, p \cdot \hat{g}$, where $\hat{g} \in L \setminus 0$.[6] ■

Preferences of the $\alpha$-MEU form are not differentiable at constant acts. If these are the only points at which $V(\cdot)$ is not differentiable (as is the case for Hurwicz preferences, defined below) then the Clarke differential is actually equal to

$$\text{co}\left\{\alpha \, \text{argmin}_{p\in\mathcal{D}} \, p \cdot f + (1-\alpha)\, \text{argmax}_{p\in\mathcal{D}} \, p \cdot f : f \in L \setminus \{0\}\right\}.$$

[6] We would like to thank the referee and associate editor for their helpful comments and suggestions in constructing the proof of this result.

If there are other points where $V(\cdot)$ is not differentiable, it is possible that $\partial V(0)$ is a proper subset of $\mathrm{co}\left\{\alpha \operatorname{argmin}_{p\in\mathcal{D}} p\cdot f + (1-\alpha)\operatorname{argmax}_{p\in\mathcal{D}} p\cdot f : f \in L\setminus\{0\}\right\}$. Whether the set inclusion is strict or not, the next result shows that extreme points of $\mathcal{D}$ are not contained in this set and hence, as an immediate corollary to Lemma 4, are not included in $\partial V(0)$.

**Lemma 5** *Let $\mathcal{D}$ be a closed convex subset of $\Delta(S)$ with cardinality greater than $1$; let $V : \mathbb{R}^n \to \mathbb{R}$ be an $\alpha$-MEU preference function with set of priors $\mathcal{D}$ and $0 < \alpha < 1$. If $\hat{p}$ is an extreme point of $\mathcal{D}$, then*

$$\hat{p} \notin \mathrm{co}\left\{\alpha \operatorname{argmin}_{p\in\mathcal{D}} p\cdot f + (1-\alpha)\operatorname{argmax}_{p\in\mathcal{D}} p\cdot f : f \in L\setminus\{0\}\right\}.$$

**Proof.** By construction, $\mathcal{D}$ has full rank within $L$. Thus if we view vectors in $L\setminus\{0\}$ as functionals on $\mathcal{D}$ none of them is constant on $\mathcal{D}$, that is, for all $f \in L\setminus\{0\}$, $\operatorname{argmin}_{p\in\mathcal{D}} p\cdot f \cap \operatorname{argmax}_{p\in\mathcal{D}} p\cdot f = \varnothing$. By definition, an extreme point of $\mathcal{D}$ cannot be written as a convex combination of two other distinct elements of $\mathcal{D}$. Therefore $\hat{p} \notin \mathrm{co}\left\{\alpha \underline{p}^f + (1-\alpha)\bar{p}^f : f \in L\setminus\{0\}\right\}$. ■

In finite dimensions, a closed convex set always contains an extreme point $\hat{p}$.[7] Thus in conjunction with Lemmas 3, 4 and 5, we have established there exists a point $\hat{p}$ in $\mathcal{D}$ such that $\hat{p} \notin \partial V(0)$. However this constitutes a failure of the preferences to admit a representation of the form given in expression (2) with the restriction that $\mathcal{D} = \partial V(0)$. Hence we have established the following result.

**Theorem 2** *Let $\mathcal{D} \subseteq \Delta(S)$ be a closed convex subset with cardinality greater than $1$; let $V : \mathbb{R}^n \to \mathbb{R}$ be an $\alpha$-MEU preference function with set of priors $\mathcal{D}$ and $0 < \alpha < 1$ and let $\succcurlyeq$ be a preference order on $\mathbb{R}^n$, which is represented by $V$. Then $\succcurlyeq$ cannot satisfy GMM axioms 1-5 and 7.*

## 3.2 Examples

We illustrate our analysis by considering two examples, Hurwicz preferences and the case where the set of priors consists of the convex combinations of two probability distributions.

### 3.2.1 Hurwicz Preferences

Hurwicz preferences are defined as follows.

[7] Indeed by the Krein Milman theorem ([5], p. 440) a closed convex set is the closure of the convex hull of its extreme points.

**Definition 5** *The* Hurwicz *preference functional,*[8] $H : A(S) \to \mathbb{R}$ *is defined by*

$$H(f) = \alpha \min_{P \in \Delta(S)} \int_S f(s)\, dP(s) + (1-\alpha) \max_{P \in \Delta(S)} \int_S f(s)\, dP(s)$$

*or equivalently* $H(f) = \alpha f_{(n)} + (1-\alpha) f_{(1)}$. *Here for a given vector* $f \in \mathbb{R}^n$ *let* $f_{(k)}$ *denotes the kth highest component of* $f$. *Hence* $f_{(1)} \geqslant f_{(2)} \geqslant ... \geqslant f_{(n)}$.

Figure 2 illustrates how our analysis applies to Hurwicz preferences when there are 3 states. In this case the set of priors is $\Delta(S)$, which has full dimension. The space $L^{\perp}$ consists just of the constant acts. Let $f$ be a given non-constant act. The dashed lines connect points at which the expected value of $f$ is constant (in probability space). For any non-constant act, the maximum and minimum expected utility occur at two distinct vertices of the simplex. For the given act $f$, the maximum and minimum expected utility occurs at $p_1 = 1$ and $p_3 = 1$ respectively. The decision weight on $f$ is therefore $\langle 1-\alpha, 0, \alpha \rangle$. In general, the decision weight, $\alpha \underline{p}^g + (1-\alpha)\bar{p}^g$, on any non-constant act, $g$, must be one of the following six vectors: $\langle \alpha, 1-\alpha, 0 \rangle$, $\langle \alpha, 0, 1-\alpha \rangle$, $\langle 1-\alpha, \alpha, 0 \rangle$, $\langle 1-\alpha, 0, \alpha \rangle$, $\langle 0, \alpha, 1-\alpha \rangle$ and $\langle 0, 1-\alpha, \alpha \rangle$. The Clarke differential $\partial H(0)$ is accordingly the convex hull of these six vectors, which forms a hexagon inside the simplex. This set is clearly closed. The extreme points of $\Delta(S)$ are the three vertices, $p_1 = 1$, $p_2 = 1$ and $p_3 = 1$. As can be seen from figure 1, for any $\alpha$, $0 < \alpha < 1$, these points are not contained in $\partial H(0)$. Moreover it is only the three vertices which are not contained in $\partial H(0)$ for all $\alpha : 0 < \alpha < 1$, i.e. for any other point in $\Delta(S)$ there is a range of values of $\alpha$ for which the given point is contained in $\partial H(0)$.[9]

[8] See Hurwicz [9] & [10]

[9] For further details of how the GMM representation applies to Hurwicz preferences see [6].

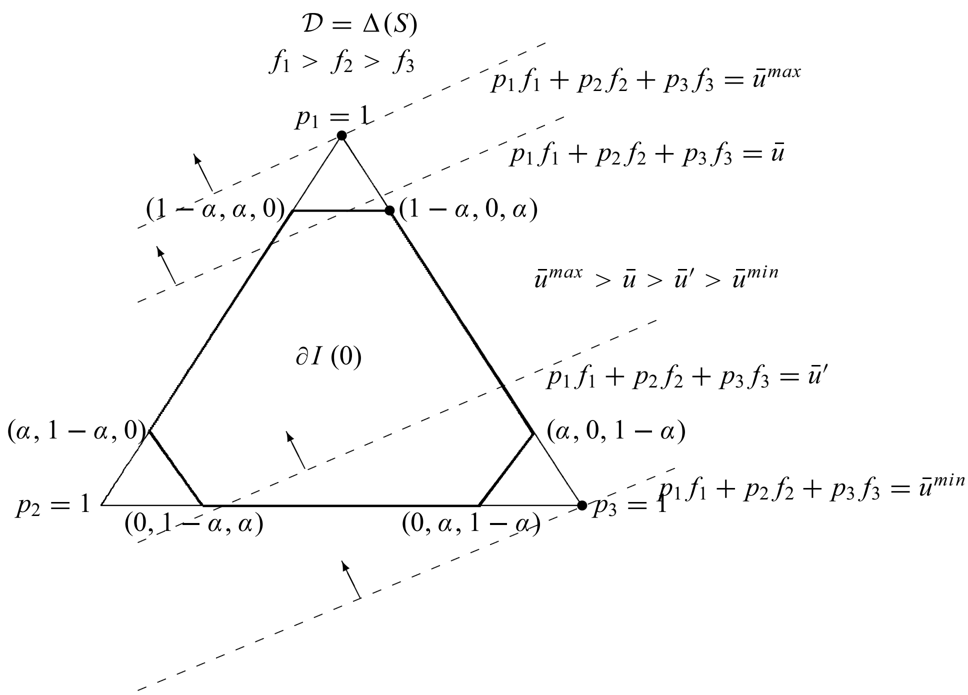



Figure 1: $\mathcal{D}$ has full rank

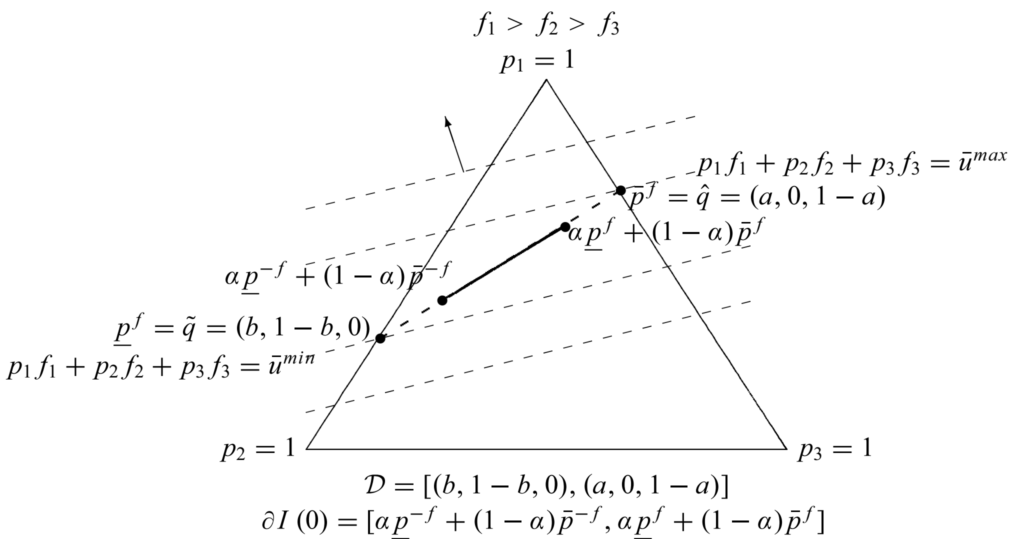



Figure 2: $\mathcal{D}$ has less than full rank

### 3.2.2 One-dimensional set of priors

In figure 3, the set of priors consists of all convex combinations of two probability distributions $\hat{q} = \langle a, 0, 1-a \rangle$ and $\tilde{q} = \langle b, 1-b, 0 \rangle$. The set of priors is a one-dimensional subset of the simplex and hence does not have full dimension. In this case $L^{\perp} = \left\{ f \in A(S) : \hat{q} \cdot f = \tilde{q} \cdot f \right\}$. This is a two dimensional subspace of $\mathbb{R}^3$, which contains the constant acts. Graphically it consists of acts whose indifference surfaces are parallel to the line connecting $\hat{q}$ and $\tilde{q}$. The given act $f$, attains its maximum at $\hat{q}$ and its minimum at $\tilde{q}$. Accordingly it gets decision weight $\alpha\tilde{q} + (1-\alpha)\hat{q}$. The Clarke differential, $\partial V(0)$, is equal to the shorter line shown in bold. In this case the extreme points are just $\hat{q}$ and $\tilde{q}$. As in the previous case, the extreme points are not contained in the Clarke differential for any value of $\alpha : 0 < \alpha < 1$. All other members of the set of priors are contained in the Clarke differential for some range of $\alpha$'s.

# 4 Infinite State Spaces

In this section we show by example that when the state space is infinite Axioms 1 - 5 and 7 can be satisfied.[10] That is, we find a set of preferences with a representation of the form given in expression (2) with an $\alpha$ in $(0, 1)$ that also satisfies the constraint $\mathcal{D} = \partial V(0)$. Indeed, our example shows it is possible to construct a set $\mathcal{D}$ which is *independent* of $\alpha$.[11]

Let the set of states of nature be $S = [0, 1]$ and let $\Sigma$ denote the $\sigma$-algebra of Borel sets of $[0, 1]$. Assume that acts lie in $C(S)$, the space of continuous functions on $S$ with the sup norm. The topological dual of $C(S)$ may be identified with $\mathrm{ca}[0, 1]$ the set of all countably additive, bounded and Borel-measurable set-functions, where the topology on $\mathrm{ca}[0, 1]$ is given by the total variation norm. If $s \in S$, let $\delta_s$ denote the Dirac measure on $S$, i.e. $\delta_s(A) = 1$, if $s \in A$; $= 0$, otherwise. Let $\mathcal{H}$ denote the set of all countably additive probability distributions on $[0, 1]$ and consider the following preference functional.

[10] In private correspondence, Klaus Nehring has informed us of an example satisfying the GMM axioms in which the set of priors is the set of all finitely additive measures on $[0, 1]$, which assign zero probability to all events of Lebesgue measure zero.

[11] It is not immediately clear from the representation in equation (3) (see page 14) that such a $\mathcal{D}$ would exist.

**Definition 6** *Define a preference functional* $W : C(S) \to \mathbb{R}$ *by*

$$W(f) = \alpha \min_{p \in \mathcal{H}} \int f dp + (1-\alpha) \max_{p \in \mathcal{H}} \int f dp.$$

*Let* $\succsim'$ *denote the preference relation on* $C(S)$ *defined by* $f \succsim' g \Leftrightarrow W(f) \geqslant W(g)$.

These preferences may be seen as the infinite dimensional analogue of the Hurwicz preferences discussed in section 3.[12] We shall show that $W(\cdot)$ satisfies the fixed point property, $H = \partial W(0)$, hence the preferences generated by $W(.)$ satisfy GMM's axiomatization.

**Proposition 2** *The preference relation* $\succsim'$ *satisfies GMM's axioms 1-5 plus 7.*

In order to prove this result we use Lemma 6 and the following two results which describe properties of the Clarke differential of a real-valued function on an arbitrary Banach space, (not necessarily $\mathbb{R}^n$).

**Proposition 3** *(Clarke [4], Corollary 2, p.39) Let* $X$ *be a Banach space and let* $V$ *and* $W$ *be real-valued functions on* $X$*. For any* $\alpha, \beta \in \mathbb{R}$, $\partial(\alpha V + \beta W)(x) \subseteq \alpha \partial V(x) + \beta \partial W(x)$.

**Proposition 4** *(Clarke [4], Proposition 2.2.7) Let* $U$ *be an open convex subset of a Banach space* $X$*. If* $V$ *is convex (resp. concave) on* $U$ *and Lipschitz near* $x$*, then* $\partial V(x)$ *coincides with the sub-gradient (resp. super-gradient) at* $x$ *in the sense of convex analysis.*

Define functionals $\xi$ (resp. $\zeta$): $A(S) \to \mathbb{R}$ by $\xi(f) = \min_{p \in \mathcal{H}} \mathbf{E}_p f$, (resp. $\zeta(f) = \max_{p \in \mathcal{H}} \mathbf{E}_p f$). Lemma 6 shows that Dirac measures are sub-gradients of $\xi$ and therefore are in the Clarke differential $\partial\xi(f)$.

**Lemma 6** *Let* $f \in C[0,1]$ *be such that* $\hat{s} \in \operatorname{argmin}_{s \in [0,1]} f(s)$, *(resp.* $\hat{s} \in \operatorname{argmax}_{s \in [0,1]} f(s)$*) then* $\delta_{\hat{s}} \in \partial\xi(f)$ *(resp.* $\delta_{\hat{s}} \in \partial\zeta(f)$*).*

**Proof.** By Proposition 4, it is sufficient to show that the linear functional $\chi : C[0,1] \to \mathbb{R}$, defined by $\chi(h) = \int h\delta_{\hat{s}}$ is a super-gradient of $\xi$ at $f$. Let $g \in C[0,1]$. Then $\xi(f) = f(\hat{s}) = \int f d\delta_{\hat{s}}$ and $\xi(g) \leqslant g(\hat{s}) = f(\hat{s}) + [g(\hat{s}) - f(\hat{s})] = \xi(f) + \int (f-g)\delta_{\hat{s}}$. This establishes that $\chi$ is a super-gradient of $\xi$ at $f$. The other case is similar. ■

[12] We would like to thank an associate editor for suggesting this argument.

**Proof of Proposition 2** As explained earlier, GMM's axioms 1-5 plus 7 are equivalent to the following representation:

$$V(f) = \beta \min_{P \in \mathcal{D}} \int_S x(s)\, dP(s) + (1-\beta) \max_{P \in \mathcal{D}} \int_S x(s)\, dP(s) \text{ and } \partial V(0) = \mathcal{D}, \tag{3}$$

for some constant $\beta$ in $[0, 1]$. We shall demonstrate that $\partial W(0) = \mathcal{H}$. The rest of the representation is clearly satisfied.

Let $\hat{s}$ be a given point in $(0, 1)$. Then there is a piecewise-linear function $f_n \in C(S)$ such that $f_n(0) = 0$, $f_n\left(\hat{s} - \frac{1}{n} - \frac{1}{n^2}\right) = 0$, $f_n\left(\hat{s} - \frac{1}{n}\right) = -\frac{1}{n}$, $f_n\left(\hat{s} - \frac{1}{n} + \frac{1}{n^2}\right) = 0$, $f_n\left(\hat{s} + \frac{1}{n} - \frac{1}{n^2}\right) = 0$, $f_n\left(\hat{s} + \frac{1}{n}\right) = \frac{1}{n}$, $f_n\left(\hat{s} + \frac{1}{n} + \frac{1}{n^2}\right) = 0$, $f_n(1) = 0$. (Thus $f_n$ is a function which has a unique maximum at $\hat{s} + \frac{1}{n}$ and a unique minimum at $\hat{s} - \frac{1}{n}$.) The function $f_n$ is illustrated in figure 4. The sequence of functions $f_n$ converges to 0, (in the sup norm).

It is clear that $\delta_{\hat{s}+\frac{1}{n}} = \operatorname{argmax}_{q \in \mathcal{H}} \int f_n dq$ and $\delta_{\hat{s}-\frac{1}{n}} = \operatorname{argmin}_{q \in \mathcal{H}} \int f_n dq$. Thus by Lemma 6 and Proposition 3, $w_n = \alpha \delta_{\hat{s}-\frac{1}{n}} + (1-\alpha)\, \delta_{\hat{s}+\frac{1}{n}} \in \partial W(f_n)$. Since $W$ is positively homogenous by [7] Proposition A.3, $\partial W(f_n) \subseteq \partial W(0)$. Hence $w_n \in \partial W(0)$. Define

$$\mathcal{J} = \overline{\text{co}} \left\{ \alpha \delta_{\hat{s}-\frac{1}{n}} + (1-\alpha)\, \delta_{\hat{s}+\frac{1}{n}} : \hat{s} \in (0,1),\, 1 \leqslant n \leqslant \infty,\, \hat{s} - \frac{1}{n} \in (0,1),\, \hat{s} + \frac{1}{n} \in (0,1) \right\},$$

where the bar denotes closure in the weak* topology. Clearly $\mathcal{J} \subseteq \mathcal{H}$.

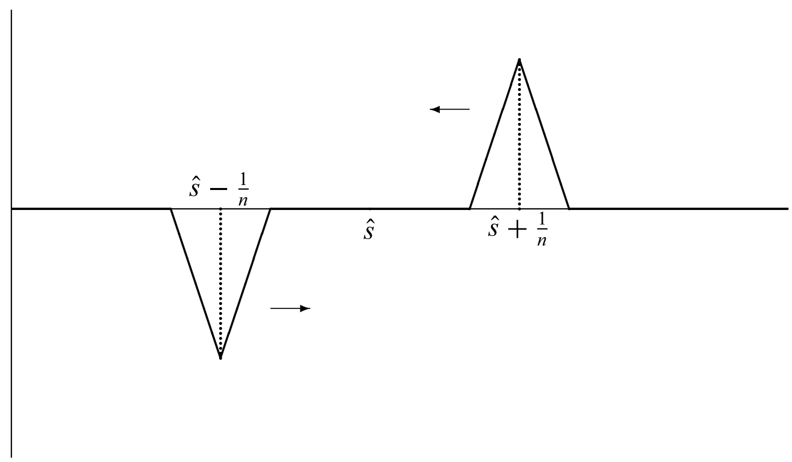



Figure 3: The function $f_n$

Since for any $g \in C(S)$, $\int g d w_n \to g(\hat{s}) = \int g d\delta_{\hat{s}}$, the sequence $w_n$ weak* converges to $\delta_{\hat{s}}$. This establishes that the Dirac measures are in $\mathcal{J}$. The convex hull of the Dirac measures is the set of discrete measures on $[0, 1]$. By Bauer [1] (Corollary 7.7.4, p. 230) the weak* closure of the discrete measures is the set of all countably additive measures on $[0, 1]$. In other words the fixed point property, $\mathcal{H} = \partial W(0)$ holds. ■

As argued above, these preferences may be seen as the infinite dimensional analogue of the Hurwicz preferences. In both cases the set of beliefs is the closed convex hull of the Dirac measures on the relevant state space. It is clear that the Dirac measures are extreme points of the set $\mathcal{H}$. For a finite state space, the set of priors is the set of all convex combinations of those probability distributions which assign probability one to a given state, that is, the Dirac measures. If there are a finite number, $n$ say, of states, then there are $n$ Dirac measures. In this case, the topology on the state space is discrete. Hence each state is topologically isolated. No state is a limit of a sequence of other states and hence the Dirac measures are not the limit of a sequence of other Dirac measures.

For the preferences studied in this section, the state space is $[0, 1]$ with the usual topology. In this case any state may be approximated by a sequence of other states and consequently any of the Dirac measures may be approximated by a sequence of other Dirac measures. If the set of priors had one or more isolated exposed points then a similar problem would arise as in finite dimensions and the GMM axioms would not be satisfied.

In infinite dimensions it is possible to construct a sequence $f_n$ such that the distance between the maximizer of $\mathbf{E} f_n$ and the minimizer of $\mathbf{E} f_n$ tends to zero as $n$ tends to infinity. This is not possible in finite dimensions, even if the set $\mathcal{D}$ has an infinite number of extreme points, since the maximizer and minimizer of $\mathbf{E} f_n$ will lie on opposite sides of the set $\mathcal{D}$ as illustrated in figure 1.

There are a number of ways in which we could extend this example. For instance assume that the state space, $S$, is any given closed convex subset of $\mathbb{R}^n$, and the space of acts is the set of continuous real valued functions on $S$. Then if $\mathcal{D}$ consists of all countably additive measures over any closed convex subset of $S$, one can show using a similar argument that the GMM axioms will be satisfied. Another interesting case is where the set of beliefs consists of convex combinations of a given prior, $q$, and an arbitrary countably additive measure, $p$, on $[0, 1]$, i.e. $\mathcal{D} = \{(1 - \gamma) q + \gamma p : p \in \mathcal{H}\}$. This can be recognized as a version of the neo-additive preferences axiomatized in Chateauneuf et al. [3]. Both of these cases can be shown to satisfy the GMM axioms by similar reasoning to that used in the proof of Proposition 2.

However even with an infinite state space, the need to satisfy a fixed-point property limits the membership of the family of preference relations which can admit a representation $V(\cdot)$ of the form in expression (2) with $\alpha$ in $(0, 1)$ and satisfying $D = \partial V(0)$. By similar reasoning to that used in section 3, the set of priors $\mathcal{D}$ cannot be finitely generated.[13] That is, $\mathcal{D}$ cannot be the set of all convex combinations of a given finite set of probability distributions. More generally these constraints cannot be satisfied when the set $\mathcal{D}$ lies in a finite dimensional (affine) subspace of $\mathrm{ca}(S)$. Another case where the GMM axioms cannot be satisfied for an $\alpha$ in $(0, 1)$ is where $\mathcal{D}$ contains an isolated extreme point. (Since the isolated extreme point will not be in the Clarke differential $\partial V(0)$.)

An open problem is to find a characterization of those sets of priors over infinite states spaces which satisfy the GMM axioms. As explained above, expression (3) imposes constraints, which imply that not any set of priors can satisfy these axioms. The precise implications of these constraints are not clear.

## Appendix: GMM Axioms 1-5 and 7.

As a reference for the reader, we list here GMM's axioms 1-5 and 7.

**Axiom 1 (Weak order)** *For all $f, g, h \in A(S)$ : (1) either $f \succsim g$ or $g \succsim f$, (2) if $f \succsim g$ and $g \succsim h$, then $f \succsim h$.*

**Axiom 2 (Certainty Independence)** *For all $f, g \in A(S)$, all $x \in X$, and all $\lambda \in (0, 1]$ :*

$$f \succsim g \Leftrightarrow \lambda f + (1-\lambda) x \succsim \lambda g + (1-\lambda) x.$$

**Axiom 3 (Archimedean Axiom)** *For all $f, g, h \in A(S)$, if $f \succ g$ and $g \succ h$, then there exist $\lambda, \mu \in (0, 1)$ such that*

$$\lambda f + (1-\lambda) h \succ g \text{ and } g \succ \mu f + (1-\mu) h.$$

**Axiom 4 (Monotonicity)** *For all $f, g \in A(S)$, if $f(s) \succsim g(s)$ for all $s \in S$, then $f \succsim g$.*

**Axiom 5 (Nondegeneracy)** *There are $f, g \in A(S)$ such that $f \succ g$.*

[13] This result has been independently proved in Siniscalchi [13].

In order to state the last axiom, recall that $\succsim^*$ is the maximal sub-relation of $\succsim$ that satisfies all the axioms of subjective expected utility except completeness.

**Axiom 7** *For all $f, g \in A(S)$, if $f \succsim^* x \Leftrightarrow g \succsim^* x$ and $x \succsim^* f \Leftrightarrow x \succsim^* g$ for all $x \in X$, then $f \sim g$.*

## References


[1] H. Bauer. *Probabilty Theory and Elements of Measure Theory*. Holt Rinehart and Winston, New York, 1972.

[2] T. Bewley. Knightian decision theory part I. *Decisions in Economics and Finance*, 2:79–110, 2002.

[3] A. Chateauneuf, J. Eichberger, and S. Grant. Choice under uncertainty with the best and worst in mind: NEO-additive capacities. *Journal of Economic Theory*, 137:538–567, 2007.

[4] F. H. Clarke. *Optimization and Nonsmooth Analysis*. Society for Industrial and Applied Mathematics, Philadelphia, 1983.

[5] N. Dunford and J. T. Schwartz. *Linear Operators*. Wiley, New York, 1958.

[6] J. Eichberger, S. Grant, and D. Kelsey. Differentiating ambiguity: An expository note. *Economic Theory*, 38:327–336, 2008.

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

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

[9] L. Hurwicz. Optimality criteria for decision making under ignorance. *Discussion paper 370, Cowles Comission*, 1951.

[10] L. Hurwicz. Some specification problems and application to econometric models. *Econometrica*, 19:343–344, 1951.

[11] K. Nehring. Imprecise probabilitistic beliefs as a context of decision-making under ambiguity. *Journal of Economic Theory*, 144:1054–1091, 2009.

[12] D. Schmeidler. Subjective probability and expected utility without additivity. *Econometrica*, 57:571–587, 1989.

[13] M. Siniscalchi. Notes on $\alpha$-MEU preferences. *private correspondence*, 2005.