
ELSEVIER
Volume 146 Number 4 July 2011 ISSN 0022-0531
JOURNAL OF Economic Theory
Editors: Christian Hellwig and Karl Shell
Associate Editors
Marco Battaglini
Dirk Bergemann
Alberto Bisin
Yeon-Koo Che
In-Koo Cho
Olivier Compte
Philip H. Dybvig
Larry G. Epstein
Yossi Feinberg
Christian Ghiglino
Jacob Goeree
Piero Gottardi
Jean-Michel Grandmont
Hugo Hopenhayn
Johannes Horner
Vijay Krishna
Felix Kubler
Ricardo Lagos
Massimo Marinacci
Thomas Mariotti
David Martimort
Kiminori Matsuyama
Richard McLean
Giuseppe Moscarini
Efe Ok
Alessandro Pavan
James Peck
Lasse Heje Pedersen
Fabrizio Perri
Benjamin Polak
Stephen A. Ross
Hamid Sabourian
William H. Sandholm
Manuel S. Santos
Chris Shannon
Marciano Siniscalchi
Lones Smith
William L. Thomson
Juuso Välimäki
Randall Wright
Administrative Editors: Edward McDevitt and Susan Witherow Schulze
Available online at
ScienceDirect
www.sciencedirect.com

Available online at www.sciencedirect.com
ScienceDirect
Journal of Economic Theory 146 (2011) 1684–1698
JOURNAL OF Economic Theory
www.elsevier.com/locate/jet

Notes

The α-MEU model: A comment ☆

Jürgen Eichberger [a], Simon Grant [b,c,∗], David Kelsey [d], Gleb A. Koshevoy [e]

[a] *Alfred Weber Institut, Universität Heidelberg, Germany*
[b] *Department of Economics, Rice University, PO Box 1892, Houston, TX 77251-1892, USA*
[c] *School of Economics, University of Queensland, Australia*
[d] *Department of Economics, University of Exeter, England, United Kingdom*
[e] *Central Institute of Mathematics and Economics RAS, Moscow, Russia*

Received 25 August 2005; final version received 27 July 2010; accepted 23 August 2010

Available online 29 March 2011

Abstract

In Ghirardato et al. (2004) [7], Ghirardato, Macheroni and Marinacci 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 α-MEU preferences. If attention is restricted to finite state spaces, we show that any α-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].

JEL classification: D81

Keywords: Ambiguity; Multiple priors; Invariant biseparable; Clarke derivative; Ambiguity-preference

☆ Research in part supported by ESRC grant No. RES-000-22-0650 and a Leverhulme Research Fellowship. 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.

∗ Corresponding author at: Department of Economics, Rice University, PO Box 1892, Houston, TX 77251-1892, USA. Fax: +1 71 3 348 5278.

E-mail addresses: juergen.eichberger@awi.uni-heidelberg.de (J. Eichberger), sgrant@rice.edu (S. Grant), D.Kelsey@exeter.ac.uk (D. Kelsey), gleb_koshevoy@mail.ru (G.A. Koshevoy).

0022-0531/$ – see front matter
doi:10.1016/j.jet.2011.03.019

1. Introduction

Ghirardato, Macheroni and Marinacci [7] (henceforth GMM), axiomatize a class of preferences which they refer to as *invariant biseparable*. This class encompasses both the Choquet expected utility (henceforth CEU) model of Schmeidler [11] and the maxmin expected utility model (also known as the multiple prior model) of Gilboa and Schmeidler [8].[1] Let $\succcurlyeq$ be an invariant biseparable preference order on the set of acts which map states from a set S to consequences in a set X. GMM define 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 $\mathcal{P}$ defined on the state space S such that for any pair of acts f and g,

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

The relation $\succcurlyeq^*$ is complete if and only if $\mathcal{P}$ 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:

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

They show that the set $\mathcal{P}$ also admits a straightforward differential characterization. Typically M 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] For our purposes it is enough to note that for the functional $I : \mathbb{R}^S \to \mathbb{R}$, which satisfies $I(u \circ f) = M(f)$, for each act f, the set $\mathcal{P}$ 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 measures $\mathcal{P}$ and an arbitrary index $\beta(\cdot)$ and substitute them into Eq. (1). Rather in order for expression (1) to generate an invariant biseparable preference relation for a given set $\mathcal{P}$ and index $\beta(.)$, we need to check that the associated Clarke differential of I at 0 is indeed equal to $\mathcal{P}$.

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

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

[1] See GMM [7, Axioms 1–5]. The formal statements of these axioms appear in Appendix A below.

[2] Note this is equivalent but not identical to the original definition, for details see [7]. Related research can be found in Nehring [10].

[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 is also a generalization of the super-gradient to a function which is not necessarily concave. For further details see Clarke [4].

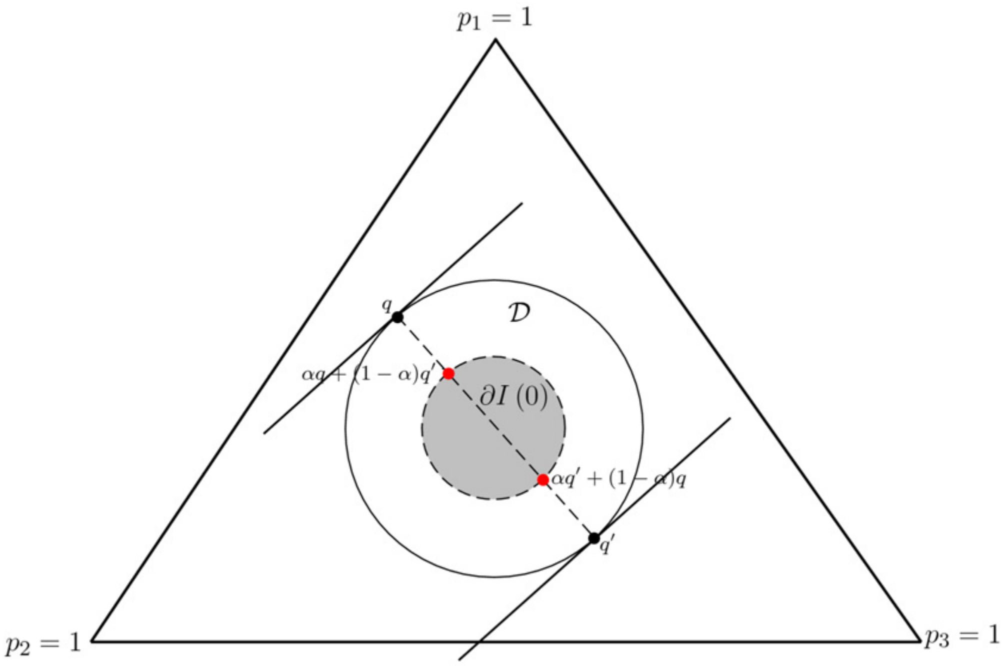

Fig. 1. Circular set of priors.

However there are two differences. First in the classic α-MEU model the weight α on the minimum expected utility is constant, whereas in expression (1) the weight $\beta(f)$ depends on the act f. Second in the classic α-MEU model the set $\mathcal{D}$ can be any non-empty weak* compact set of probability measures, whereas in expression (1), $\mathcal{P}$ 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 α 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 $\beta(f)$ 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 α-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 for any relation that satisfies GMM's axiomatization (that is, their Axioms 1–5 and 7) the constant ambiguity-aversion index, α, is equal either to 0 or to 1. Or equivalently, the preference relation is either maxmax expected utility or maxmin expected utility.

Our strategy of proof is to fix a closed convex set, $\mathcal{D}$, of probability distributions on a finite set S and an α 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-aversion index, α, 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 Fig. 1. The figure considers a given act f. The expected value of f is maximised at $q \in \mathcal{D}$ and minimised at q'.

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

The probability used to calculate the expectation of f (and hence value $V(f)$) is accordingly $\alpha q' + (1-\alpha)q$. As f varies, the corresponding probability associated with each act traces out the boundary of the inner circle. The Clarke differential, $\partial I(0)$, is the convex hull of these points, which is represented by 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 α-MEU preferences satisfying GMM's axioms in this case.

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 α-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 Appendix A.

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. *(See 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, ∇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 continuously 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. Next we define the Clarke differential, which 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

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

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 continuously 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. *(See 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*

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

where Γ_V denotes the set of points at which V is differentiable and $\mathrm{co}(A)$ *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 α 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, if we take invariant biseparable preferences represented by expression (1) and impose the extra restriction that $\beta(f)$ be a constant function then a fixed point property must be satisfied. We show that a fixed point only exists if $\beta(f) \equiv 1$ or $\beta(f) \equiv 0$. In particular, if $\beta(f) \equiv \alpha$ for some α 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 $\phi(f) = \min_{p\in\mathcal{D}} p \cdot f$ and $\psi(f) = \max_{p\in\mathcal{D}} p \cdot f$. That is, ϕ and ψ represent maxmin and maxmax expected utility preferences respectively. The functions ϕ and ψ 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.[5] However, since ϕ is concave and ψ 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 with radius ϵ 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 ϕ is differentiable, the minimizing probability distribution is unique and is equal to the derivative. It also finds an expression for the derivative of V at points where both ϕ and ψ 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. $\operatorname{argmax}_{p\in\mathcal{D}} p \cdot f$).

[5] By the dimension of $\mathcal{D}$ we mean the dimension of the affine space spanned by $\mathcal{D}$.

Lemma 3.

1. *If ϕ (resp. ψ) is differentiable at f then* $\operatorname{argmin}_{p\in\mathcal{D}} p.f$ *(resp.* $\operatorname{argmax}_{p\in\mathcal{D}} p.f$*) is unique.*
2. *Suppose that ϕ (resp. ψ) is differentiable at f, then $d\phi_f(y) = \underline{p}^f \cdot y$ (resp. $d\psi_f(y) = \bar{p}^f \cdot y$) for all $y \in \mathbb{R}^n$. This may be expressed in terms of gradients as $\underline{p}^f = \nabla\phi(f)$ (resp. $\bar{p}^f = \nabla\psi(f)$).*
3. *Let V be an α-MEU preference functional. If ϕ and ψ 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 parts 1 and 2 for ϕ. A similar argument applies to ψ. Suppose that ϕ is differentiable at f. Then since $d\phi_f$ is a linear function on $\mathbb{R}^n$, there exists $z \in \mathbb{R}^n$ such that $d\phi_f(y) = z.y$, for all $y \in \mathbb{R}^n$. Let $\underline{p}^f$ be an element of $\operatorname{argmin}_{p\in\mathcal{D}} p.f$. Suppose, if possible, $z \neq \underline{p}^f$. Consider $h \in \mathbb{R}^n$ such that $\underline{p}^f.h = 0$ and $z.h > 0$. Let $\tilde{q}$ be an element of $\operatorname{argmin}_{p\in\mathcal{D}} p.(f+\epsilon h)$. Then $\tilde{q}.(f+\epsilon h) \leqslant \underline{p}^f.(f+\epsilon h)$. Thus $\frac{\phi(f+\epsilon h)-\phi(f)-d\phi_f(\epsilon h)}{\|\epsilon h\|} = \frac{\tilde{q}.(f+\epsilon h)-\underline{p}^f.f-\epsilon z.h}{\epsilon\|h\|} \leqslant \frac{\underline{p}^f.(f+\epsilon h)-\underline{p}^f.f-\epsilon z.h}{\epsilon\|h\|} = \frac{\underline{p}^f.f+\epsilon\underline{p}^f.h-\underline{p}^f.f-\epsilon z.h}{\epsilon\|h\|} = -\frac{\epsilon z.h}{\epsilon\|h\|} = -\frac{z.h}{\|h\|} < 0$. Hence $\lim_{\epsilon\to 0}\frac{\phi(f+\epsilon h)-\phi(f)-d\phi_f(\epsilon h)}{\|\epsilon h\|} \leqslant -\frac{z.h}{\|h\|} < 0$. However this contradicts the assumption that ϕ is differentiable at f. Thus we may conclude that $z = \underline{p}^f$. Parts 1 and 2 of the lemma now follow. Part 3 follows from part 2 and linearity of the derivative on $\mathbb{R}^n$. □

Let L denote the linear span of $\{p-q\colon\ p, q \in \mathcal{D}\}$ and denote by $L^\perp$ the orthogonal complement of the vector space L.[6] If $\mathcal{D}$ has full rank then $L^\perp$ will consist just of the constant vectors in $A(S) = \mathbb{R}^n$.[7] If the dimension of $\mathcal{D}$ is less than $n-1$, then $L^\perp$ 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.

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

$$\partial V(0) \subseteq \operatorname{co}\bigl\{\alpha \operatorname{argmin}_{p\in\mathcal{D}} p\cdot f + (1-\alpha)\operatorname{argmax}_{p\in\mathcal{D}} p\cdot f\colon\ f \in L\backslash\{0\}\bigr\}.$$

Proof. Since ϕ is a concave function and ψ 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,

$$\begin{aligned}\partial V(0) &= \operatorname{co}\bigl\{\lim \nabla V(f_n)\colon\ f_n \to 0,\ f_n \in \Gamma_V \cap \Gamma_\phi \cap \Gamma_\psi\bigr\}\\ &= \operatorname{co}\bigl\{\lim\bigl(\alpha\underline{p}^{f_n} + (1-\alpha)\bar{p}^{f_n}\bigr)\colon\ f_n \to 0,\ f_n \in \Gamma_V \cap \Gamma_\phi \cap \Gamma_\psi\bigr\}.\end{aligned}$$

Note that $L^\perp \cap (\Gamma_\phi \cap \Gamma_\psi) = \emptyset$, because at $h \in L^\perp$, ϕ and ψ are not differentiable. This follows, provided $\mathcal{D}$ is not a singleton, since $h \in L^\perp$ implies $p\cdot h = p'\cdot h$ for all $p, p' \in \mathcal{D}$ and hence, $\operatorname{argmin}_{p\in\mathcal{D}} p\cdot h$ and $\operatorname{argmax}_{p\in\mathcal{D}} p\cdot h$ are not singletons.

---

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

7 We say that $\mathcal{D}$ has full rank if the dimension of L is $n-1$.

Consider a particular sequence $f_n \to 0$, $f_n \in \Gamma_V \cap \Gamma_\phi \cap \Gamma_\psi$ such that $\lim(\alpha \underline{p}^{f_n} + (1-\alpha)\bar{p}^{f_n})$ exists. Fix an n. Write $f_n = g_n + h_n$, where $g_n \in L$ and $h_n \in L^\perp$. Since $f_n \in \Gamma_\phi \cap \Gamma_\psi$ we know that $f_n \notin L^\perp$ and hence $g_n \neq 0$. Define $\hat{g}_n = \frac{g_n}{\|g_n\|}$. 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 $\bar{p} \in \operatorname{argmax}_{p \in \mathcal{D}} p \cdot \hat{g}$ and $\underline{p} \in \operatorname{argmin}_{p \in \mathcal{D}} p \cdot \hat{g}$. Putting the steps together we have $\lim(\alpha \underline{p}^{f_n} + (1-\alpha)\bar{p}^{f_n}) = \alpha \underline{p} + (1-\alpha)\bar{p} \in \alpha \operatorname{argmin}_{p \in \mathcal{D}} p \cdot \hat{g} + (1-\alpha) \operatorname{argmax}_{p \in \mathcal{D}} p \cdot \hat{g}$, where $\hat{g} \in L \backslash 0$.[8] □

Preferences of the α-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

$$\operatorname{co}\{\alpha \operatorname{argmin}_{p \in \mathcal{D}} p \cdot f + (1-\alpha) \operatorname{argmax}_{p \in \mathcal{D}} p \cdot f\colon\ f \in L \backslash \{0\}\}.$$

If there are other points where $V(\cdot)$ is not differentiable, it is possible that $\partial V(0)$ is a proper subset of $\operatorname{co}\{\alpha \operatorname{argmin}_{p \in \mathcal{D}} p \cdot f + (1-\alpha) \operatorname{argmax}_{p \in \mathcal{D}} p \cdot f\colon\ f \in L \backslash \{0\}\}$. 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 α-MEU preference function with set of priors $\mathcal{D}$ and $0 < \alpha < 1$. If $\hat{p}$ is an extreme point of $\mathcal{D}$, then*

$$\hat{p} \notin \operatorname{co}\{\alpha \operatorname{argmin}_{p \in \mathcal{D}} p \cdot f + (1-\alpha) \operatorname{argmax}_{p \in \mathcal{D}} p \cdot f\colon\ f \in L \backslash \{0\}\}.$$

Proof. By construction, the affine span of $\mathcal{D}$ is a translation of the subspace L. Thus if we view vectors in $L \backslash \{0\}$ as functionals on $\mathcal{D}$ none of them is constant on $\mathcal{D}$, that is, for all $f \in L \backslash \{0\}$, $\operatorname{argmin}_{p \in \mathcal{D}} p \cdot f \cap \operatorname{argmax}_{p \in \mathcal{D}} p \cdot f = \emptyset$. By definition, an extreme point of $\mathcal{D}$ cannot be written as a convex combination of two other distinct elements of $\mathcal{D}$. Therefore $\hat{p} \notin \operatorname{co}\{\alpha \underline{p}^f + (1-\alpha)\bar{p}^f\colon\ f \in L \backslash \{0\}\}$. □

In finite dimensions, a closed convex set always contains an extreme point $\hat{p}$.[9] 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.

---

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

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

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 α-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.

Definition 5. The *Hurwicz* preference functional,[10] $H : A(S) \to \mathbb{R}$ 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 $f \in \mathbb{R}^n$, $f_{(k)}$ denotes the kth highest component of f. Hence $f_{(1)} \geqslant f_{(2)} \geqslant \cdots \geqslant f_{(n)}$.

Fig. 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 probability used in evaluating the expectation of f is therefore $\langle 1 - \alpha, 0, \alpha \rangle$. In general, the probability, $\alpha \underline{p}^g + (1 - \alpha)\bar{p}^g$, used to evaluate the expectation of 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 Fig. 2, for any α, $0 < \alpha < 1$, 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 α: $0 < \alpha < 1$, i.e. for any other point in $\Delta(S)$ there is a range of values of α for which the given point is contained in $\partial H(0)$.[11]

3.2.2. One-dimensional set of priors

In Fig. 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 $L^{\perp} = \{f \in A(S)\colon \hat{q} \cdot f = \tilde{q} \cdot f\}$. 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,

[10] See Hurwicz [9]. A more detailed discussion of these preferences can be found in L. Hurwicz. Optimality criteria for decision making under ignorance. Discussion paper 370, Cowles Commission, 1951.

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

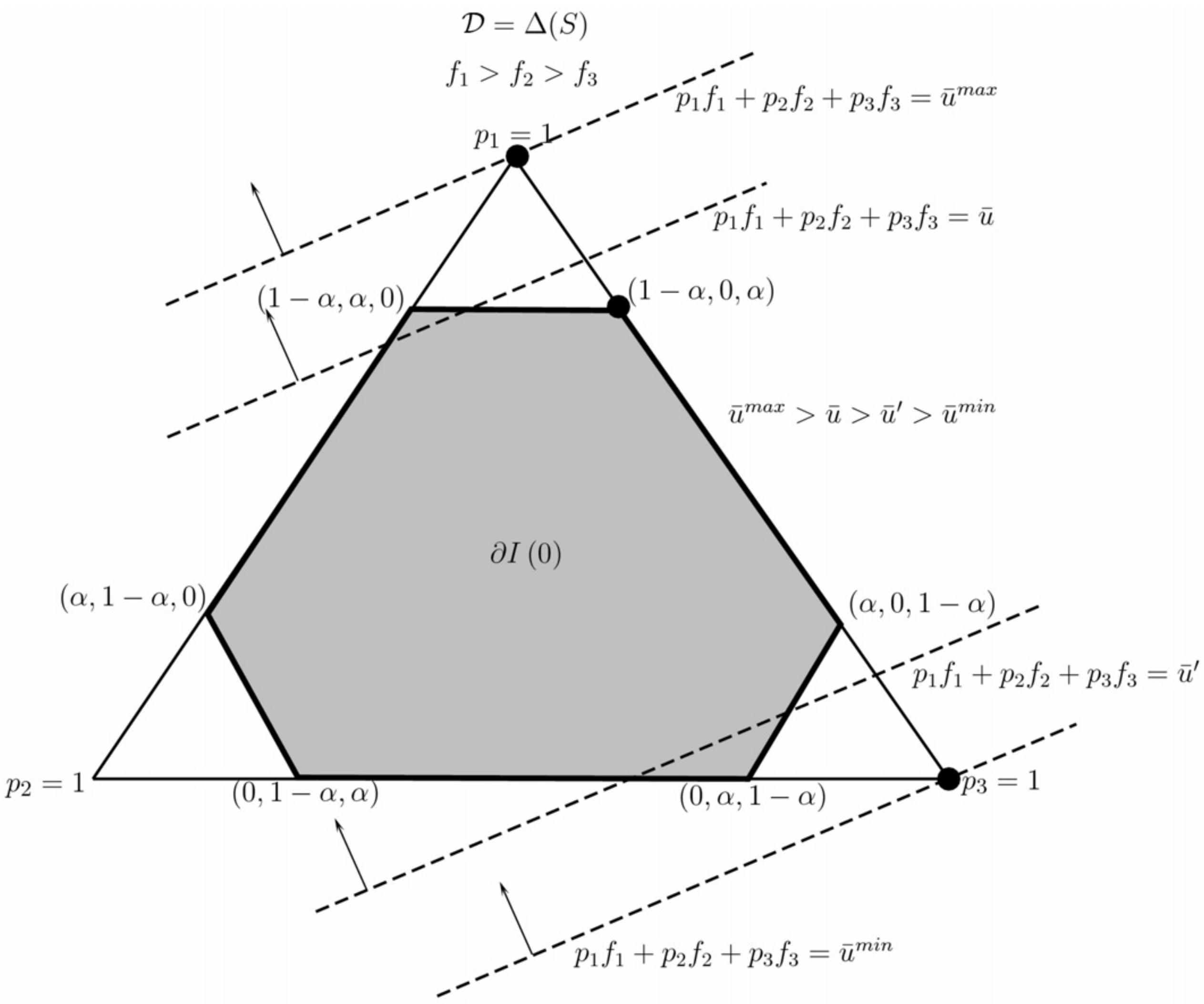

Fig. 2. $\mathcal{D}$ has full rank.

attains its maximum at $\hat{q}$ and its minimum at $\tilde{q}$. Accordingly its expectation is taken with respect to the probability $\alpha\tilde{q} + (1-\alpha)\hat{q}$. 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 α: $0 < \alpha < 1$. All other members of the set of priors are contained in the Clarke differential for some range of α'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.[12] That is, we find a set of preferences with a representation of the form given in expression (2) with an α 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 α.[13]

[12] 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.

[13] It is not immediately clear from the representation in Eq. (3) (see p. 1695) that such a $\mathcal{D}$ would exist.

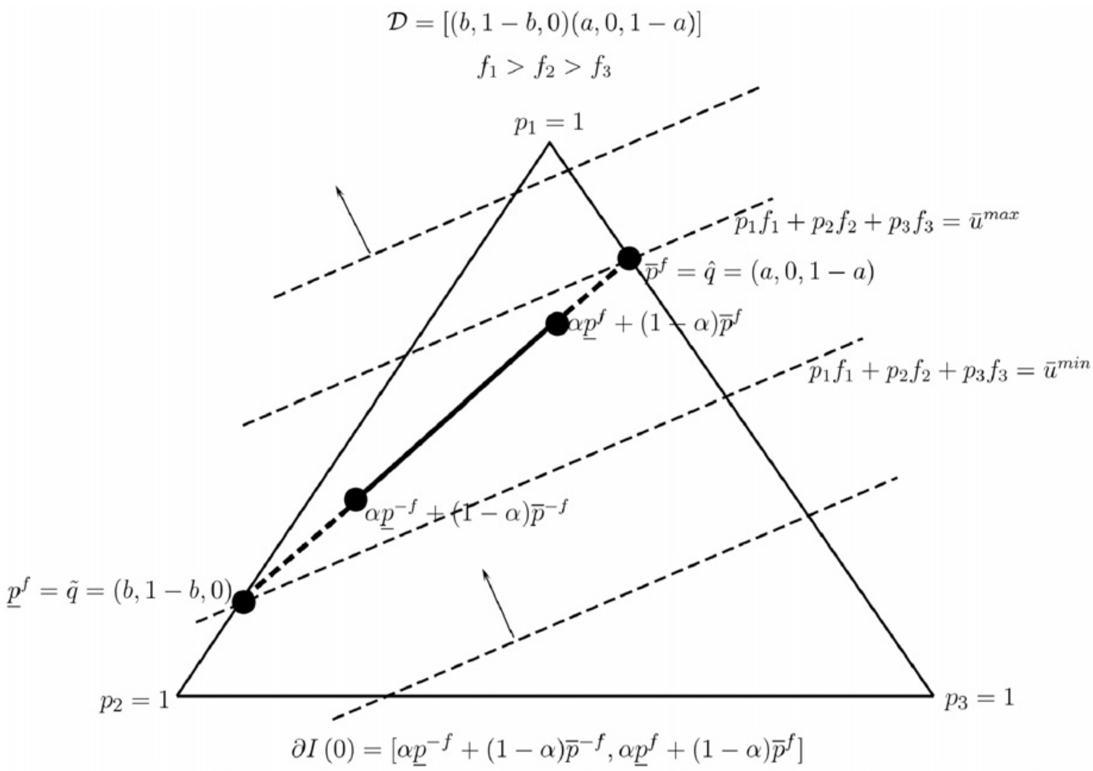

Fig. 3. $\mathcal{D}$ has less than full rank.

Let the set of states of nature be $S = [0, 1]$ and let Σ denote the σ-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 ca[0, 1] the set of all countably additive, bounded and Borel-measurable set-functions, where the topology on ca[0, 1] is given by the total variation norm. If $s \in S$, let δ_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.

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

These preferences may be seen as the infinite-dimensional analogue of the Hurwicz preferences discussed in Section 3.[14] We shall show that $W(\cdot)$ satisfies the fixed point property, $\mathcal{H} = \partial W(0)$, hence the preferences generated by $W(.)$ satisfy GMM's axiomatization.

Proposition 2. *The preference relation $\succcurlyeq'$ 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 $\mathbb{R}^n$).

[14] We would like to thank the associate editor for suggesting this argument.

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

Proposition 4. *(See 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 ξ (resp. ζ) : $A(S) \to \mathbb{R}$ by $\xi(f) = \min_{p\in\mathcal{H}} \int f\,dp$ (resp. $\zeta(f) = \max_{p\in\mathcal{H}} \int f\,dp$). Lemma 6 shows that if $\hat{s}$ minimizes $f \in C[0,1]$, then the Dirac measure $\delta_{\hat{s}}$ is a super-gradient of ξ at f and therefore is 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. $\tilde{s} \in \operatorname{argmax}_{s\in[0,1]} f(s)$) then $\delta_{\hat{s}} \in \partial\xi(f)$ (resp. $\delta_{\tilde{s}} \in \partial\zeta(f)$).*

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 ξ at f. Let $g \in C[0,1]$. Then $\xi(f) = f(\hat{s}) = \int f\,d\delta_{\hat{s}}$ and $\xi(g) \leqslant g(\hat{s}) = f(\hat{s}) + [g(\hat{s}) - f(\hat{s})] = \xi(f) + \int(g - f)\delta_{\hat{s}}$. This establishes that χ is a super-gradient of ξ at f. The other case is similar. □

Proof of Proposition 2. As explained earlier, GMM's Axioms 1–5 plus 7 are equivalent to the following representation:

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

for some constant β in [0, 1]. We shall demonstrate that for $W(f)$ from Definition 6 one obtains $\partial W(0) = \mathcal{H}$. The rest of the representation is clearly satisfied.

Let $\hat{s}$ be a given point in (0, 1). Fix an integer $n > 0$. Then there is a piecewise-linear function $f_n \in C(S)$ such that $f_n(0) = 0$, $f_n(\hat{s} - \frac{1}{n} - \frac{1}{n^2}) = 0$, $f_n(\hat{s} - \frac{1}{n}) = -\frac{1}{n}$, $f_n(\hat{s} - \frac{1}{n} + \frac{1}{n^2}) = 0$, $f_n(\hat{s} + \frac{1}{n} - \frac{1}{n^2}) = 0$, $f_n(\hat{s} + \frac{1}{n}) = \frac{1}{n}$, $f_n(\hat{s} + \frac{1}{n} + \frac{1}{n^2}) = 0$, $f_n(1) = 0$. (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 Fig. 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

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

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

Since for any $g \in C(S)$, $\int g\,dw_n \to g(\hat{s}) = \int g\,d\delta_{\hat{s}}$, 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 noted 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

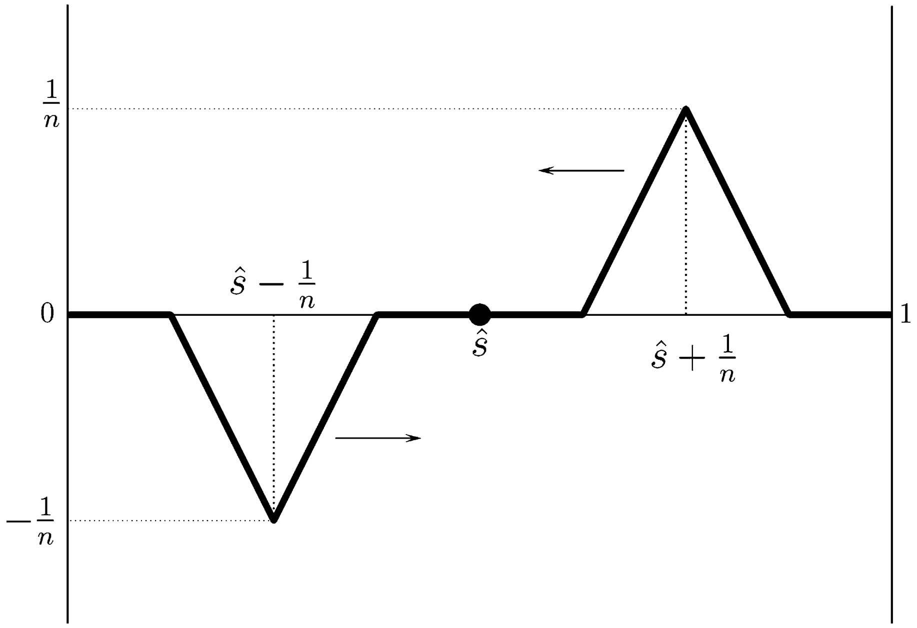

Fig. 4. The function f_n.

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 extreme 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 $\operatorname{argmin}_{q \in \mathcal{H}} \int f_n \, dq$ and $\operatorname{argmax}_{q \in \mathcal{H}} \int f_n \, dq$ are unique and the maximizer of $\int f_n \, dq$ and the minimizer of $\int f_n \, dq$ converge to a common limit 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 $\int f_n \, dq$ will lie on opposite sides of the set $\mathcal{D}$ as illustrated in Fig. 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. $\mathcal{D} = \{(1 - \gamma)q + \gamma p\colon\ p \in \mathcal{H}\}$. 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 α in $(0, 1)$ and satisfying $\mathcal{D} = \partial V(0)$. By similar reasoning to that used in Section 3, the set of priors $\mathcal{D}$ cannot be finitely generated.[15] 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 α 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 A. 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 $\lambda \in (0,1]$:

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

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

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

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 $f \succ g$.

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, Probability Theory and Elements of Measure Theory, Holt Rinehart and Winston, New York, 1972.
[2] T. Bewley, Knightian decision theory part I, Decis. Econ. Finance 2 (2002) 79–110.
[3] A. Chateauneuf, J. Eichberger, S. Grant, Choice under uncertainty with the best and worst in mind: NEO-additive capacities, J. Econ. Theory 137 (2007) 538–567.
[4] F.H. Clarke, Optimization and Nonsmooth Analysis, SIAM Publ., Philadelphia, 1983.

[15] In private correspondence Marciano Siniscalchi has informed us that he has an independent proof of this result.

[5] N. Dunford, J.T. Schwartz, Linear Operators, Wiley, New York, 1958.
[6] J. Eichberger, S. Grant, D. Kelsey, Differentiating ambiguity: An expository note, Econ. Theory 38 (2008) 327–336.
[7] P. Ghirardato, F. Macheroni, M. Marinacci, Differentiating ambiguity and ambiguity attitude, J. Econ. Theory 118 (2004) 133–173.
[8] I. Gilboa, D. Schmeidler, Maxmin expected utility with a non-unique prior, J. Math. Econ. 18 (1989) 141–153.
[9] L. Hurwicz, Some specification problems and application to econometric models, Econometrica 19 (1951) 343–344.
[10] K. Nehring, Imprecise probabilistic beliefs as a context of decision-making under ambiguity, J. Econ. Theory 144 (2009) 1054–1091.
[11] D. Schmeidler, Subjective probability and expected utility without additivity, Econometrica 57 (1989) 571–587.