Eliciting beliefs

Robert Chambers · Tigran Melkonyan

Abstract   We develop an algorithm that can be used to approximate a decisionmaker's beliefs for a class of preference structures that includes, among others, α-maximin expected utility preferences, Choquet expected utility preferences, and, more generally, constant additive preferences. For both exact and statistical approximation, we demonstrate convergence in an appropriate sense to the true belief structure.

Keywords α-Maximin expected utility · Choquet expected utility · Rank-dependent · Invariant biseparable preferences · Constant additive preferences · Geometric tomography

The elicitation of subjective probabilities, via the betting odds system devised by de Finetti (1937, 1972), is conceptually central to the theory of decisionmaking under uncertainty. Procedures for eliciting subjective probabilities are available in both the state-independent, subjective expected utility framework (e.g., Savage 1971; Davis and Holt 1993) and the state-dependent, subjective expected utility framework (Karni 1999). However, repeated empirical validations of both the Allais Paradox and the Ellsberg Paradox have forced the realization that subjective expected utility theory may not accommodate reality. As a consequence, a variety of competing preference representations have been developed (Kahneman and Tversky 1979; Quiggin 1981, 1982; Bewley 1986, 1987; Schmeidler 1989; Gilboa and Schmeidler 1989; Tversky and Kahneman 1992; Epstein and Zhang 2001; Ghirardato and Marinacci 2001; among many others).

R. Chambers
Agricultural and Resource Economics, University of Maryland, College Park, 2200 Symons Hall, College Park, MD 20817, USA

T. Melkonyan (✉)
Resource Economics, University of Nevada, Reno, MS 204, Reno, NV 89557, USA
e-mail: tmelkonyan@cabnr.unr.edu

This article develops an algorithm that can be used to approximate an individual's belief structure for a class of decision theoretic models that includes many of these more general preference structures. Our algorithm also provides an approximation to the class of imprecise probability models (Walley 1991) where information is represented by a set of probability measures. This model is in one-to-one correspondence with Walley's (1991) model of coherent lower previsions which, according to Walley (2000), is the most general theory of imprecise probability in situations not involving conditioning on events with probability zero.

The elicitation procedure is based upon the generalization of de Finetti's (1937, 1972) betting-odds system that underlies the modern theory of imprecise probabilities (Smith 1961; Levi 1980; Good 1983; Walley 1991). Both exact approximation and statistical approximation are considered. Exact approximation refers to the case where an individual's willingness to pay can be measured without sampling error, while statistical approximation refers to the case where willingness to pay is measured with sampling error. We provide an upper bound on the measurement error (in terms of the Hausdorff distance between the true belief structure and the constructed belief structure) in both instances, and we demonstrate that both procedures converge in an appropriate sense to the true belief structure.

The approach relies on the recognition that, for a very broad class of models, an individual's marginal willingness to sell or marginal willingness to buy a stochastic asset can be represented mathematically as a support functional for the individual's belief structure (in a sense made precise below).[1] This recognition allows the exact and statistical approximation algorithms that have emerged in the literatures on geometric tomography, pattern recognition, and signal processing (e.g., Gardner 1995) to be applied to the belief elicitation problem in a simple and straightforward manner.

There is a large and growing literature on belief elicitation in models that allow for violations of the axioms of expected utility theory. Most studies take rank-dependent utility (RDU) theory or cumulative prospect theory (CPT) as starting points and then estimate decision weights for the Bernoulli utility functions that are embedded in these structures (see, for example, Cohen and Jaffray (1988), Currim and Sarin (1989), Starmer and Sugden (1989), Camerer and Ho (1994), Mangelsdorff and Weber (1994), Fox and Tversky (1998), Wu and Gonzalez (1999), Viscusi and Chesson (1999), Abdellaoui (2000), Bleichrodt and Pinto (2000), Kilka and Weber (2001), Diecidue et al. (2003), Abdellaoui et al. (2005)).

Our elicitation procedure is for a class of preferences that includes α-maximin expected utility, Choquet expected utility (CEU), rank-dependent, invariant biseparable and, more generally, constant additive preferences, and smooth variational preferences. By introducing a reference point and considering both gains and losses our approach can be utilized to elicit beliefs of a decisionmaker with CPT preferences.

The approach that we take does not impose any specific structure on the underlying preference function (such as additive separability over co-monotonic acts). Instead, it imposes structure on a differential representation of preferences. Specifically, we treat

---

[1] A number of others have used local approximations to study decision-making under uncertainty, an approach owing much of its popularity to Machina (1982).

preference structures for which beliefs are proportional to an appropriately normalized Clarke differential of the preference functional (Clarke 1983).

We emphasize, however, that what we term "beliefs" may encompass behavioral traits other than "true beliefs" about the stochastic environment (see Ghirardato et al. 2004, 2005 for a discussion and example). As our procedure is defined for a number of different models, the interpretation of the results obtained using that procedure necessarily varies across those models. Moreover, the interpretation of the empirically elicited "beliefs" will also depend on the specifics of the experimental design used to elicit them and upon the maintained theoretical model of the "true belief" structure. In contrast to the "relation-based" axiomatic approach where there is no knowledge about the decisionmaker's information, a properly designed experiment, within the context of a maintained model of the underlying belief structure, can attempt to filter out various behavioral traits of "beliefs," such as perception of ambiguity and probabilistic risk aversion, by controlling the experimental subjects' information on the stochastic environment. Thus, the economic interpretation of the 'beliefs' elicited using our procedure depends critically upon the empirical setting and upon the maintained model of the true beliefs.

In what follows, we first define our notation and make some preliminary observations. Then we specify the class of preference structures. After that, we present two algorithms (one exact and one statistical) to approximate the underlying belief structure. The article then concludes.

1 Notation and preliminaries

Ω denotes a finite state space, $\Omega = \{1, \ldots, S\}$. Random variables are defined as maps from the state space to $\Re$, the set of reals. Σ denotes the σ-algebra of all subsets of Ω. Δ denotes the set of all additive probability measures over Ω. A generic element of Δ is denoted by $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_S)$, where π_s is the probability of elementary event $\{s\}$.

Random variables can be identified with vectors in $\Re^{\Omega}$, which is endowed with the Euclidean norm $|\cdot|$ and inner product. The degenerate random variable is denoted by $\mathbf{1} = (1, \ldots, 1) \in \Re^{\Omega}$. For $\mathbf{x}, \mathbf{y} \in \Re^{\Omega}$, we represent the inner product by $\mathbf{x}'\mathbf{y} \equiv \sum_{s=1}^{S} x_s y_s$. The unit sphere is denoted $M \subset \Re^{\Omega}$ where

$$M = \left\{ \mathbf{x} \in \Re^{\Omega} : |\mathbf{x}| = 1 \right\}.$$

The (upper) *support function* $h(K, \cdot)$ for a non-empty closed convex set $K \subset \Re^{\Omega}$ is defined by

$$h(K, \mathbf{u}) \equiv \sup \left\{ \mathbf{k}'\mathbf{u} : \mathbf{k} \in K \right\} \quad \text{for } \mathbf{u} \in \Re^{\Omega},$$

and the (lower) support function is

$$-h(K, -\mathbf{u}) = -\sup \left\{ \mathbf{k}'(-\mathbf{u}) : \mathbf{k} \in K \right\} \quad \text{for } \mathbf{u} \in \Re^{\Omega}$$
$$= \inf \left\{ \mathbf{k}'\mathbf{u} : \mathbf{k} \in K \right\} \quad \text{for } \mathbf{u} \in \Re^{\Omega}.$$

The Hausdorff distance between two sets $K, L \subset \Re^{\Omega}$ is defined by

$$\delta(K, L) = \max \left\{ \sup_{\mathbf{k} \in K} \inf_{\mathbf{l} \in L} |\mathbf{k} - \mathbf{l}|, \sup_{\mathbf{l} \in L} \inf_{\mathbf{k} \in K} |\mathbf{k} - \mathbf{l}| \right\} \text{ for } K, L \subseteq \Re^{\Omega}. \quad (1)$$

For convex compact sets $K, L \subset \Re^{\Omega}$ (Schneider 1993):

$$\delta(K, L) \equiv \sup_{\mathbf{u} \in M} \left\{ |h(K, \mathbf{u}) - h(L, \mathbf{u})| \right\}.$$

2 Preferences

Preferences over random variables are given by the functional, $W : \Re^{\Omega} \to \Re$. For our elicitation procedure to work, preferences must satisfy a minimum smoothness requirement. W is assumed to be Lipschitzian (of rank N) in the neighborhoods of the degenerate random variables. That is, for any $\lambda \in \Re$ there exists $\varepsilon > 0$ such that $|W(\mathbf{y}^1) - W(\mathbf{y}^2)| \leq N |\mathbf{y}^1 - \mathbf{y}^2|$ for all $\mathbf{y}^1, \mathbf{y}^2 \in (\lambda \mathbf{1} + \varepsilon B)$. Ghirardato et al. (2004) demonstrate that preferences satisfying the axioms underlying the class of invariant biseparable preferences satisfy this smoothness requirement.

When W is Lipschitzian (of some rank N) near $\mathbf{y}$, it has a well-defined (lower) generalized directional derivative at $\mathbf{y}$ in the direction of $\mathbf{v}$ defined by

$$W'(\mathbf{y}; \mathbf{v}) = \liminf_{\mathbf{x} \to \mathbf{y}, t \downarrow 0} \left\{ \frac{W(\mathbf{x} + t\mathbf{v}) - W(\mathbf{x})}{t} \right\}.$$

$W'(\mathbf{y}; \mathbf{v})$ is finite and superlinear (positively linearly homogeneous and superadditive) in $\mathbf{v}$ (Clarke 1983). As it is superlinear, it is the (lower) support function for a closed, convex set. That closed, convex set is the generalized gradient (Clarke differential)[2] and is given by

$$\partial W(\mathbf{y}) = \left\{ \boldsymbol{\pi} \in \Re^{\Omega} : \boldsymbol{\pi}'\mathbf{v} \geq W'(\mathbf{y}; \mathbf{v}) \text{ for all } \mathbf{v} \in \Re^{\Omega} \right\}. \quad (2)$$

From Proposition 2.1.2 in Clarke (1983)

$$\begin{aligned} W'(\mathbf{y}; \mathbf{v}) &= \inf \left\{ \boldsymbol{\pi}'\mathbf{v} : \boldsymbol{\pi} \in \partial W(\mathbf{y}) \right\} \\ &= -h(\partial W(\mathbf{y}), -\mathbf{v}), \end{aligned} \quad (3)$$

and

$$-W'(\mathbf{y}; -\mathbf{v}) = h(\partial W(\mathbf{y}), \mathbf{v}). \quad (4)$$

These directional derivatives are interpreted, respectively, as the marginal benefit in the neighborhood of $\mathbf{y}$ accruing to the individual from acquiring control of the random variable, $\mathbf{v}$, and losing control of the random variable $\mathbf{y}$.

---

[2] The terminology Clarke differential is due to Ghirardato and Marinacci (2004).

Our remaining assumption on the preference structure is satisfied by a very broad class of state-independent preference structures (expected utility, rank-dependent utility, α-maximin expected utility, invariant biseparable preferences, and, more generally, constant additive preferences).

Assumption For any $\lambda \in \mathfrak{R}$, there exists $\mu > 0$ such that

$$\partial W (\lambda \mathbf{1}) = \mu \mathcal{P}, \tag{5}$$

where $\mathcal{P} \subset \Delta$ is closed and convex.

Broadly speaking, the assumption ensures that in the neighborhood of the degenerate random variables, preferences are decomposable into two components. One (μ) measures the decisionmaker's tastes and the other ($\mathcal{P}$) represents his or her beliefs about the state of the world. Operationally, the assumption is important because it implies that the individual's marginal willingness to sell (MWS) an asset $\mathbf{v}$ as measured in units of the degenerate random variable is

$$\frac{-W' (\lambda \mathbf{1}; -\mathbf{v})}{W' (\lambda \mathbf{1}; \mathbf{1})} = \frac{h (\mu \mathcal{P}, \mathbf{v})}{-h (\mu \mathcal{P}, -\mathbf{1})}$$
$$= h (\mathcal{P}, \mathbf{v}),$$

because $h (\mu \mathcal{P}, \mathbf{v}) = \mu h (\mathcal{P}, \mathbf{v})$ and $-h (\mu \mathcal{P}, -\mathbf{1}) = \mu$.

We close this section with some examples of classes of preferences that satisfy (5).

Example Let preferences be of the variational class derived by Maccheroni et al. (2006)

$$W (\mathbf{y}) = \min \left\{ \boldsymbol{\pi}' \mathbf{u} (\mathbf{y}) + c (\boldsymbol{\pi}) : \boldsymbol{\pi} \in \boldsymbol{\Delta} \right\},$$

where $c (\boldsymbol{\pi})$ is convex and lower semicontinuous, $\mathbf{u} (\mathbf{y}) = [u (y_1), \dots, u (y_S)]$, and u is the decisionmaker's concave and smooth utility of deterministic outcomes. Then

$$\partial W (\lambda \mathbf{1}) = u' (\lambda) \arg \min \{u (\lambda) + c (\boldsymbol{\pi}) : \boldsymbol{\pi} \in \boldsymbol{\Delta}\}$$
$$= u' (\lambda) \arg \min \{c (\boldsymbol{\pi}) : \boldsymbol{\pi} \in \boldsymbol{\Delta}\}.$$

Example More generally, let

$$W (\mathbf{y}) \equiv V (\mathbf{u} (\mathbf{y})),$$

where V is nondecreasing in all its arguments. Constant additive preferences satisfy

$$V (\mathbf{u} + \boldsymbol{\alpha}) = V (\mathbf{u}) + \alpha,$$

where $\mathbf{u} \in \mathfrak{R}^{\Omega}$, and $\boldsymbol{\alpha} = (\alpha, \alpha, \dots, \alpha)$ corresponds to a constant act. For the class of constant additive preferences, calculation reveals

$$V' \left(\mathbf{u}, \hat{u} + \mathbf{1} \right) = V' \left(\mathbf{u}, \hat{u} \right) + 1,$$

and

$$V'\left(\mathbf{u}, \hat{\mathbf{u}} - \mathbf{1}\right) = V'\left(\mathbf{u}, \hat{\mathbf{u}}\right) - 1,$$

so that any $\mathbf{p} \in \partial V(\mathbf{u})$ must satisfy

$$\mathbf{p}'\mathbf{1} \geq 1 \geq \mathbf{p}'\mathbf{1},$$

and thus $\mathbf{p} \in \Delta$. Thus, constant additive preferences with concave u satisfy (5). Special cases of constant additive preferences include invariant-biseparable preferences, CEU preferences, $\alpha - MEU$ preferences, MEU preferences, RDU preferences, and SEU preferences.

3 Algorithm to approximate $\mathcal{P}$

3.1 Approximation without noise

In this section, we assume that it is possible to obtain exact measurements of MWS without any sampling error. Denote the $S - 1$-dimensional unit sphere by

$$T = \left\{\mathbf{x} \in \mathfrak{R}^{\Omega} : x_S = 0, |\mathbf{x}| = 1\right\}.$$

For $0 < \varepsilon < 1$, $U \subseteq T$ is called an ε- *net* if for every $\mathbf{x} \in T$ there exists $\mathbf{z} \in U$ such that $|\mathbf{x} - \mathbf{z}| \leq \varepsilon$, i.e., each point in T is within an ε distance of some point in U. Due to the requirement that probabilities have to sum to 1, the maximum dimensionality of any set of probabilities is $S - 1$. This allows us to consider $S - 1$-dimensional ε-nets when the number of states of nature is S.

Figure 1 depicts an ε-net in the three-dimensional case; $S - 1 = 3$. We have not depicted dimension x_4 in Fig. 1 since $x_4 = 0$ for all points in an ε-net. The points on the sphere of radius 1 represent the endpoints of the vectors that form the ε-net. Each such vector emanates from the origin, has unitary length, and is within an ε distance of some vector in the ε-net.

Choose a set of directions $(\mathbf{y}^1, \ldots, \mathbf{y}^N)$, with each $\mathbf{y}^i \in T$, that forms an ε-net for some $\varepsilon \in (0, 1)$. Elicit from a respondent, whose initial position is $\lambda\mathbf{1}$, his or her MWS and let the report be denoted by ϕ^i. Performing this procedure for each element of $(\mathbf{y}^1, \ldots, \mathbf{y}^N)$ gives N reports of MWS. Let
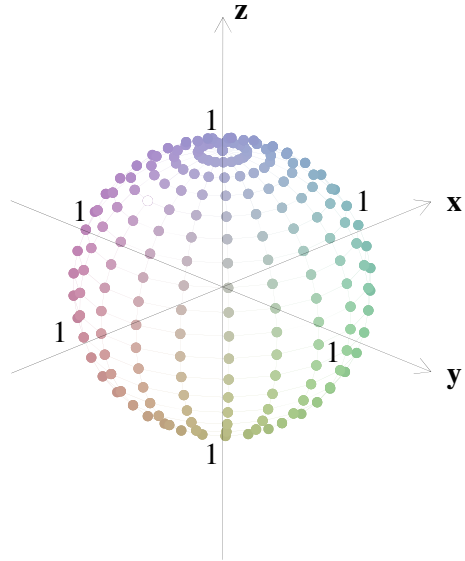
$$\widehat{\mathcal{P}} \equiv \left\{\boldsymbol{\pi} : \boldsymbol{\pi}'\mathbf{y}^i \leq \phi^i \text{ for all } i = 1, \ldots, N\right\}. \tag{6}$$

Then:

Theorem 1 *(i) $\widehat{\mathcal{P}}$ is a polytope; (ii) $h(\mathcal{P}, \mathbf{y}^i) = h(\widehat{\mathcal{P}}, \mathbf{y}^i)$ for all $i = 1, \ldots, N$; (iii) $\mathcal{P} \subseteq \widehat{\mathcal{P}}$; (iv) $\delta\left(\mathcal{P}, \widehat{\mathcal{P}}\right) \leq \frac{2\varepsilon}{1-\varepsilon}$.*

Proof (i) $\widehat{\mathcal{P}}$ is a polytope if and only if it is a bounded polyhedral set (Theorem 9.2 in Brondsted 1983). By construction, $\widehat{\mathcal{P}}$ is the intersection of a finite number of closed

Fig. 1 ε − net

halfspaces, i.e., $\widehat{\mathcal{P}}$ is a polyhedral set. $\widehat{\mathcal{P}}$ is bounded since $(\mathbf{y}^1, \ldots, \mathbf{y}^N)$ forms an ε -net; (ii) $h(\mathcal{P}, \mathbf{y}^i) = \phi^i = h(\widehat{\mathcal{P}}, \mathbf{y}^i)$ by construction. (iii) follows immediately from our construction and (ii); (iv) Lemma 4.1 in Gardner and Milanfar (2003) states that, for $\alpha > 0$ and $0 < \varepsilon < 1$, if K and L are compact convex sets in $\Re^{\Omega}$ such that $h(K, \mathbf{y}) \equiv h(L, \mathbf{y}) \leq \alpha$ for each $\mathbf{u}$ in some ε-net $U \subseteq T$ then $\delta(K, L) \leq \frac{2\varepsilon\alpha}{1-\varepsilon}$. $\mathcal{P}$ and $\widehat{\mathcal{P}}$ are compact convex sets with $h(\mathcal{P}, \mathbf{y}^i) \equiv h(\widehat{\mathcal{P}}, \mathbf{y}^i)$ for all $\mathbf{y}^i$ in an ε-net $(\mathbf{y}_1, \ldots, \mathbf{y}_N)$ (parts (i) and (ii)). Furthermore, $h(\mathcal{P}, \mathbf{y}^i) = h(\widehat{\mathcal{P}}, \mathbf{y}^i) = \sup\left\{\sum_{s=1}^{S} \pi_s y_s^i : \boldsymbol{\pi} \in \mathcal{P}\right\} \leq 1$ since $\boldsymbol{\pi} \in \mathcal{P}$ and $\left|\mathbf{y}^i\right| = 1$. □

Figure 2 provides a graphical demonstration of an approximation of $\mathcal{P}$ in a Marschak-Machina triangle. In Fig. 2, $\mathbf{y}^i$ is a representative direction that is an element of some ε-net, ϕ^i is the reported MWS for that direction and $\pi_1 y_1^i + \pi_2 y_2^i = \phi^i$ is the supporting hyperplane of $\mathcal{P}$ in the direction $-\mathbf{y}^i = (-y_1^i, -y_2^i, 0)$. $\widehat{\mathcal{P}}$ is given by polytope ABCDEF which is derived by taking the intersection of half-spaces defined by supporting hyperplanes to $\mathcal{P}$ in directions constituting the ε-net. It is clear from Fig. 2 that as ε gets smaller the number of elements $\mathbf{y}^i$ of an ε-net increases and, accordingly, one gets a smoother and closer approximation of $\mathcal{P}$.

Operationally, one can envision our elicitation approach in the following fashion. Starting from a position of certainty, one first specifies "assets" or lottery tickets that conform to each of the N directions. One then uses an elicitation procedure to determine empirically the amount at which the respondent is willing to sell a "small" amount of each of those assets. After this is done, these reports are used to construct $\widehat{\mathcal{P}}$.

Expression (6) and Theorem 1 together show that the algorithm we propose can be used to create an outer approximation to $\mathcal{P}$ that can be made to have an arbitrarily small degree of approximation error by the definition of U. It follows immediately
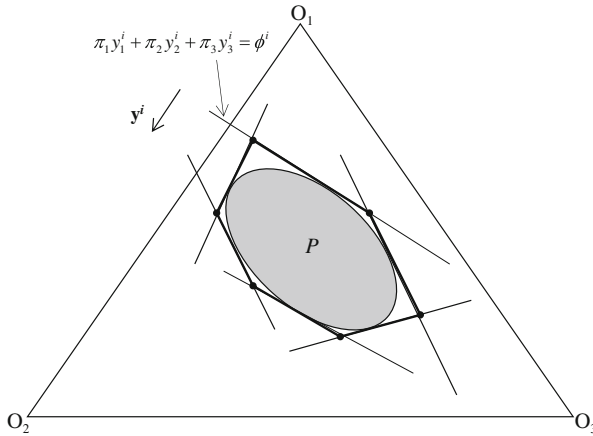
Fig. 2 Approximation of P

that as U approaches T this approximation converges in the Hausdorff distance. We state this fact as a corollary:

Corollary 2 $\lim\limits_{\varepsilon \to 0} \widehat{\mathcal{P}} = \mathcal{P}$.

By basic results from convex geometry, to minimize the computational complexity of the algorithm for given ε, as reflected by the number N of elements in the ε-net, one needs to choose elements $(\mathbf{y}^1, \ldots, \mathbf{y}^N)$ of the ε-net to form vertices of a regular polytope with distances between adjacent vertices equal to ε.[3] Denoting the number of elements in such a regular polytope by $N(\varepsilon)$, it is straightforward to verify that $\lim\limits_{\varepsilon \to 0} N(\varepsilon) = \infty$. Thus, the number of MWS reports required to get an arbitrarily close approximation to the actual set of beliefs approaches infinity.

However, when there is prior information on the underlying preferences, in certain cases, one can elicit beliefs in a finite number of steps. For example, when $\mathcal{P}^*$, obtained from $\mathcal{P}$ by

$$\mathcal{P}^* = \left\{ \boldsymbol{\rho} : A \in \Sigma, \rho(A) = \sum_{s \in A} \pi_s, (\pi_1, \pi_2, \ldots, \pi_S) \in \mathcal{P} \right\}, \tag{7}$$

is a core of a supermodular capacity, Chambers and Melkonyan (2005) observe that exposed faces of $\mathcal{P}$ are parallel to the faces of the probability simplex. Selecting directions $\mathbf{y}^i$ that exhaust the set of faces of the probability simplex ($2^S - 2$ reports of MWS), they obtain an exact representation of $\mathcal{P}$ in a finite number of iterations. Closely related to the algorithm of Chambers and Melkonyan (2005) is the observation of Dow and Werlang (1992) that the divergence between MWS and the marginal willingness to pay for an asset in the CEU model with a supermodular capacity measures the ambiguity of the individual's belief.

---

[3] For methods to construct such polytopes and their properties, the reader is referred to Brondsted (1983).

More generally, when $\mathcal{P}$ is a polytope with known directions of its faces, one can elicit beliefs in a finite number of steps by selecting directions $\mathbf{y}^i$ that correspond to the faces of the polytope. As an example, consider the case where $\mathcal{P}^*$, obtained from $\mathcal{P}$ by (7), is a core of a simple capacity. A capacity, f, is *simple* if

$$f(\{s\}) = (1 - \lambda)g(\{s\}) \text{ for all } s \in \Omega,$$
$$f(\{i_1, \ldots, i_k\}) = \sum_{j=1}^{k} f(\{i_j\}) \text{ when } k < S,$$
$$f(\Omega) = 1,$$

where $\mathbf{g} \equiv (g(\{1\}), \ldots, g(\{S\})) \in \Delta$ and $\lambda \in [0, 1]$. A simple capacity is supermodular, and

$$\mathcal{P} = \{(1 - \lambda)\mathbf{g} + \lambda\Delta\}. \tag{8}$$

$\mathcal{P}$ is, thus, the Minkowski sum of a singleton set $\{(1-\lambda)\mathbf{g}\}$ and the probability simplex Δ scaled by λ. $\mathcal{P}$ is often referred to as λ-contaminated set of probabilities. Thus, $\mathcal{P}$ is completely characterized by S parameters $(\lambda, g(\{1\}), \ldots, g(\{S - 1\}))$. Moreover, $\mathcal{P}$ is a simplex with faces that are parallel to those of simplex Δ. Choosing as directions of movement $(\mathbf{y}^1, \ldots, \mathbf{y}^S)$ directions of the faces of probability simplex Δ, one can elicit $\mathcal{P}$, or equivalently parameters $(\lambda, g(\{1\}), \ldots, g(\{S - 1\}))$, in S iterations.[4]

An alternative and equivalent approach to elicit parameters $(\lambda, g(\{1\}), \ldots, g(\{S - 1\}))$ is to first elicit the 'degree of confidence' λ and then the anchor probability $\mathbf{g}$. For the former, elicit from the decisionmaker her MWP and MWS an arbitrarily chosen asset $\mathbf{v} \equiv \{v_1, \ldots, v_S\}$ with $\max\{v_1, \ldots, v_S\} > \min\{v_1, \ldots, v_S\}$. Denoting these reports by $MWP(\mathbf{v})$ and $MWS(\mathbf{v})$ and using (3), (4) and (8), we obtain

$$MWP(\mathbf{v}) = \inf\{\boldsymbol{\pi}'\mathbf{v} : \boldsymbol{\pi} \in \{(1 - \lambda)\mathbf{g}\} + \lambda\Delta\}\} = (1 - \lambda)\mathbf{g}'\mathbf{v} + \lambda \min\{v_1, \ldots, v_S\}$$

and

$$MWS(\mathbf{v}) = \sup\{\boldsymbol{\pi}'\mathbf{v} : \boldsymbol{\pi} \in \{(1 - \lambda)\mathbf{g}\} + \lambda\Delta\}\} = (1 - \lambda)\mathbf{g}'\mathbf{v} + \lambda \max\{v_1, \ldots, v_S\}.$$

It follows immediately from the above two expressions that

$$\lambda = \frac{MWS(\mathbf{v}) - MWP(\mathbf{v})}{\max\{v_1, \ldots, v_S\} - \min\{v_1, \ldots, v_S\}}.$$

Given the knowledge of λ, one can determine the anchor probability $\mathbf{g}$ by eliciting the decisionmaker's MWP (or MWS) for $S - 1$ elementary lotteries, i.e., assets with one of the components equal to 1 and the rest equal to 0.

---

[4] We thank Chris Chambers for making this observation.

Fig. 3 Illustration of
inconsistent support function
measurements
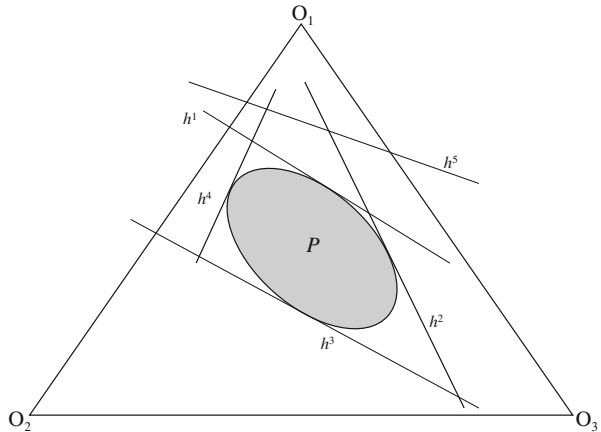
3.2 Approximation with noise

In many laboratory and natural experiments, MWS's (or MWP's) reports are presumably observed with noise that can be attributed purely to sampling error and that does not emerge from the characteristics of the underlying preferences. In the presence of such exogenous sampling error, reports of MWS may not be *consistent* in the sense that there might be no convex set $\mathcal{P}$ that could have *all* the MWS reports as support function measurements. Figure 3 provides a demonstration of this possibility. There, support function measurements h^i (for $i = 1, \ldots, 4$) are mutually consistent, but h^5 is not. That is, there does not exist a convex set that has h^i (for $i = 1, \ldots, 5$) as support function measurements.

The possibility of inconsistent observations on MWS leads naturally to the examination of conditions that ensure consistency of MWS as an observation on a support function. Karl et al. (1995) provide necessary and sufficient conditions for consistency of support function measurements. Some new terminology is necessary.

$K \subseteq \mathfrak{R}^\Omega$ is a *cone* if it is closed under positive scalar multiplication, i.e., $\lambda K \subseteq K$ for any $\lambda > 0$. $C\left(h^1, h^2, \ldots, h^n\right) = \left\{ \sum_{i=1}^n \lambda_i h^i : \lambda_i \geq 0 \text{ for } i \in \{1, 2, \ldots, n\}, \lambda_i > 0 \text{ for at least one } i \right\}$ is called a *positive cone* of the vectors $\left(h^1, h^2, \ldots, h^n\right)$. A cone is *full* if it is not contained in a proper subspace.

Theorem 3 (Karl, Kulkarni, Verghese, and Willsky) *The set of support function measurements $\left(h^1, h^2, \ldots, h^N\right)$ for a set of directions $(\mathbf{y}^1, \ldots, \mathbf{y}^N)$, that form an ε-net for some $\varepsilon \in (0, 1)$, is consistent if and only if the following determinantal condition is satisfied for all $(S + 1)-$ tuples of vectors $\left(\mathbf{y}^{i_1}, \ldots, \mathbf{y}^{i_{S+1}}\right)$ with one belonging to the full positive cone of the others:*

$$\begin{vmatrix} h^{i_1} & \left(\mathbf{y}^{i_1}\right)' \\ h^{i_2} & \left(\mathbf{y}^{i_2}\right)' \\ \vdots & \vdots \\ h^{i_{S+1}} & \left(\mathbf{y}^{i_{S+1}}\right)' \end{vmatrix} \begin{vmatrix} 1 & \left(\mathbf{y}^{i_1}\right)' \\ 1 & \left(\mathbf{y}^{i_2}\right)' \\ \vdots & \vdots \\ 1 & \left(\mathbf{y}^{i_{S+1}}\right)' \end{vmatrix} \geq 0. \tag{9}$$

Condition (9) is a requirement that support function measurement h^i (for all $i = 1, \ldots, N$) passes through the intersection of the half-spaces defined by all other support function measurements h^j for $j \neq i$. In the 2-dimensional case depicted in Fig. 2, one such requirement is that h^5 passes through the polytope ABCD.

For the relatively small dimensional problems that characterize many experimental settings, S is reasonably small (frequently only two or three). In such cases, verification of (9) is straightforward. However, in more complicated settings, verification of (9) can be computationally burdensome. Karl et al. (1995) demonstrate that condition (9) needs to be verified only for $(S + 1)-$tuples of vectors $\left(\mathbf{y}^{i_1}, \ldots, \mathbf{y}^{i_{S+1}}\right)$ that form appropriately defined "neighborhoods". This local test for consistency is considerably more efficient than the global test in Theorem 3. In order to economize on space and notation, we refer the reader to Karl et al. (1995) for details.

Suppose that the MWS reports $\left(\phi^1, \phi^2, \ldots, \phi^N\right)$ for an ε-net of directions $(\mathbf{y}^1, \ldots, \mathbf{y}^N)$ satisfy

$$\phi^i = h(\mathcal{P}, \mathbf{y}^i) + \theta^i \quad \text{for } i = 1, 2, \ldots, N,$$

where $\theta^i \sim N\left(0, \sigma^2\right)$. Our objective is to construct a set $\widehat{\mathcal{P}}$ that "best approximates", in a statistical sense, the unknown set of beliefs $\mathcal{P}$. Following Gardner et al. (2006) consider the following constrained least squares problem:

$$\min_{h^1, \ldots, h^N} \sum_{i=1}^{N} \left(\phi^i - h^i\right)^2 \tag{10}$$

subject to $h^1, \ldots, h^N$ are consistent.

Let $\widehat{h}^1, \ldots, \widehat{h}^N$ denote the solution to (10) and let

$$\widehat{\mathcal{P}}\left(\widehat{h}^1, \ldots, \widehat{h}^N\right) \equiv \left\{\boldsymbol{\pi} : \boldsymbol{\pi}'\mathbf{y}^i \leq \widehat{h}^i \text{ for all } i = 1, \ldots, N\right\}.$$

It follows immediately from Gardner et al. (2006) that

Theorem 4 *If*

$$\phi^i = h(\mathcal{P}, \mathbf{y}^i) + \theta^i \quad \text{for } i = 1, 2, \ldots, N,$$

with $\theta^i \sim N\left(0, \sigma^2\right)$, *then almost surely*

$$\lim_{N \to \infty} \delta\left(\mathcal{P}, \widehat{\mathcal{P}}\left(\widehat{h}^1, \ldots, \widehat{h}^N\right)\right) = 0.$$

A similar result holds for more general probability distributions of θ and alternative distance functions (Gardner et al. 2006).

4 Conclusion

We have used convex analysis in combination with results from geometric tomography, pattern recognition, and signal processing to develop algorithms for exact and statistical approximation of beliefs in a broad class of state-independent preference structures. Both exact and statistical approximation algorithms utilize reports of MWS (or MWP) for an appropriately chosen set of stochastic assets.

When reports of MWS are observed without error, the MWS reports correspond to the support hyperplanes to the actual set of probability distributions. Since any convex set can be represented by the set of its supporting hyperplanes (Minkowski's theorem) one can obtain an arbitrary close polytopial approximation to the actual set of probabilities by eliciting MWS's for a "sufficiently dense" set of directions.

In the presence of noise, the "best" statistical approximation will depend, among other things, on the distribution of the noise term. When the error term has a normal distribution, we consider the least squares estimator constrained by a consistency condition. The latter is a requirement that the set of fitted hyperplanes constitute supporting hyperplanes for *some* convex set. The resulting polytopial approximation converges almost surely to the actual set of probabilities. This result is robust against more general assumptions on the probability distribution.

Our analysis has been restricted to finite dimensional state spaces. This has been done to permit us to apply existing results from the geometric tomography and image processing literatures, which is largely focused on identifying bodies in finite dimensional spaces, directly to our setting. While it limits the overall theoretical generality of our results, we note that, as a practical matter, most empirical elicitation experiments are focused on relatively low dimension finite state spaces. Our results are particularly appropriate to that setting.

Although we have worked solely with approximations in terms of the Hausdorff distance function, equivalent results hold for a number of other distance functions (see Gardner et al. 2006). By using a Hausdorff metric, we are requiring that the approximation and the actual beliefs be "close" in all directions. This is very important for our application because otherwise (if for some directions the distance between the two sets is not controlled) the predicted willingness to pay for an asset (willingness to sell) and the actual willingness to pay for that asset (willingness to sell) can be arbitrarily large. Moreover, as Gardner (1995) notes, the Hausdorff metric is the "default metric" and "standard one in the study of convex bodies." The Hausdorff metric is standard in many areas including the theory of approximations, pattern recognition, robotics, and a multitude of other disciplines. The Hausdorff metric is equivalent to a number of other popular metrics when these are defined on the set of convex bodies. For example, the Hausdorff metric is equivalent to the symmetric difference metric on the set of convex bodies.

We note in closing that support functions are observed in many other areas of operations research and economics. Thus, the general methods used here should prove useful in other applications and other fields. For example, a competitive firm's profit is the support function for the convex free disposal hull of its technology (e.g., Chambers 1988). It follows directly that analogues of the methods used here can be used to approximate that hull of the technology.

References

Abdellaoui, M. (2000). A parameter-free elicitation of utility and probability weighting functions. *Management Science, 46*(11), 1497–1512.

Abdellaoui, M., Vossmann, F., & Weber, M. (2005). Choice-based elicitation and decomposition of decision weights for gains and losses under uncertainty. Forthcoming in *Management Science*.

Bewley, T. F. (1986). *Knightian decision theory: Part I.* Discussion Paper, Cowles Foundation.

Bewley, T. F. (1987). *Knightian decision theory: Part II*. Discussion Paper, Cowles Foundation.

Bleichrodt, H., & Pinto, J. L. (2000). A parameter-free elicitation of the probability weighting function in medical decision analysis. *Management Science, 46*(11), 1485–1496.

Brondsted, A. (1983). *An introduction to convex polytopes.* Springer-Verlag.

Camerer, C. F., & Ho, T.-H. (1994). Valuations of the betweenness axiom and nonlinearity in probability. *Journal of Risk and Uncertainty, 8*, 167–196.

Chambers, R. G. (1988). *Applied production analysis*. Cambridge University Press.

Chambers, R. G., & Melkonyan, T. A. (2005). Eliciting the core of a supermodular capacity. *Economic Theory, 26*(1), 203–210.

Clarke, F. H. (1983). *Optimization and nonsmooth analysis.* New York: Wiley.

Cohen, M., & Jaffray, J.-Y. (1988). Certainty effect versus probability distortion: An experimental analysis of decision making under risk. *Journal of Experimental Psychology: Human Perception and Performance, 14*(4), 554–560.

Currim, I., & Sarin, R. K. (1989). Prospect versus utility. *Management Science, 35*, 22–41.

Davis, D. D., & Holt, C. A. (1993). *Experimental economics.* Princeton: Princeton University Press.

de Finetti, B. (1937). La Prévision: Ses Lois Logiques, ses Sources Subjectives. Annales de l'Institute Henri Poincaré, 7, 1–68. Translated into English by Henry E. Kyburg Jr., "Foresight: Its logical laws, its subjective sources," In Kyburg Jr., H. E., & Smokler, H. E. (Eds.). (1964) *Studies in subjective probability* (2nd edn, pp. 53–118). New York: Wiley, 1980, New York: Krieger.

de Finetti, B. (1972). *Probability, induction and statistics. The art of guessing.* London-New York-Sydney: John Wiley & Sons.

Diecidue, E., Wakker, P. P., & Zeelenberg, M. (2003). Using prospect theory to measure beliefs in stock performances for which probabilities are unknown. Working Paper.

Dow, J., & Werlang, S. (1992). Uncertainty aversion, risk aversion, and the optimal choice of portfolio. *Econometrica*, *60*, 197–204.

Epstein, L. G., & Zhang, J. (2001). Subjective probabilities on subjectively unambiguous events. *Econometrica, 69*, 256–306.

Fox, C. R., & Tversky, A. (1998). A belief-based account of decision under uncertainty. *Management Science, 44*, 879–895.

Gardner, R. J. (1995). *Geometric tomography*. Cambridge University Press.

Gardner, R. J., Kiderlen, M., & Milanfar, P. (2006). Convergence of algorithms for reconstructing convex bodies and directional measures. *The Annals of Statistics, 34*, 1331–1374.

Gardner, R. J., & Milanfar, P. (2003). Reconstruction of convex bodies from brightness functions. *Discrete and Computational Geometry, 29*, 279–303.

Ghirardato, P., & Marinacci, M. (2001). Risk, ambiguity, and the separation of utility and beliefs. *Mathematics of Operations Research, 26*(4), 864–890.

Ghirardato, P., Maccheroni, F., & Marinacci, M. (2004). Differentiating ambiguity and ambiguity attitude. *Journal of Economic Theory, 118*(2), 133–73.

Ghirardato, P., Maccheroni, F., & Marinacci, M. (2005). Certainty independence and the separation of utility and beliefs. *Journal of Economic Theory, 120*, 129–136.

Gilboa, I., & Schmeidler, D. (1989). Maxmin expected utility with non-unique Prior. *Journal of Mathematical Economics, 18*, 141–153.

Good, I. J. (1983). *Good thinking: The foundations of probability and its applications.* Minneapolis: University of Minnesota Press.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*(2), 263–291.

Karni, E. (1999). Elicitation of subjective probabilities when preferences are state-dependent. *International Economic Review, 40*, 479–486.

Karl, W. C., Kulkarni, S. R., Verghese, G. C., & Willsky, A. S. (1995). Local tests for consistency of support hyperplane data. *Journal of Mathematical Imaging and Vision, 6*, 249–267.

Kilka, M., & Weber, M. (2001). What determines the shape of the probability weighting function under uncertainty? *Management Science, 47*(12), 1712–1726.

Levi, I. (1980). *The enterprise of knowledge.* Massachusetts: MIT Press, Cambridge.

Maccheroni, F., Marinacci, M., & Rustichini, A. (2006). Ambiguity aversion, robustness, and the variational representation of preferences. *Econometrica, 74*, 1447–1498.

Machina, M. J. (1982). 'Expected utility' Analysis without the independence axiom. *Econometrica, 50*, 277–323.

Mangelsdorff, L., & Weber, M. (1994). Testing choquet expected utility. *Journal of Economic Behavior and Organization, 25*, 437–457.

Quiggin, J. (1981). Risk perception and risk aversion among Australian farmers. *Australian Journal of Agricultural Economics, 25*, 160–169.

Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior and Organization, 3*, 323–343.

Savage, L. J. (1971). Elicitation of personal probabilities and expectations. *Journal of the American Statistical Association, 66*, 783–801.

Schmeidler, D. (1989). Subjective probability and expected utility without additivity. *Econometrica, 57*, 571–587.

Schneider, R. (1993). *Convex bodies: The Brunn-Minkowski theory.* Cambridge University Press.

Smith, C. A. B. (1961). Consistency in statistical inference and decision. *Journal of Royal Statistical Society B, 23*, 1–25.

Starmer, C., & Sugden, R. (1989). Probability and juxtaposition effects: An experimental investigation of the common ratio effect. *Journal of Risk and Uncertainty, 2*, 159–178.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: cumulative representation of uncertainty. *Journal of Risk and Uncertainty, 5*(4), 297–323.

Viscusi, K. W., & Chesson, H. (1999). Hopes and fears: The conflicting effects of risk ambiguity. *Theory and Decision, 47*, 153–178.

Walley, P. (1991). *Statistical reasoning with imprecise probabilities.* London: Chapman & Hall.

Walley, P. (2000). Towards a unified theory of imprecise probability. *International Journal of Imprecise Reasoning, 24*, 125–148.

Wu, G., & Gonzalez, R. (1999). Nonlinear decision weights in choice under uncertainty. *Management Science, 45*, 74–85.