Econ Theory
DOI 10.1007/s00199-009-0503-8

RESEARCH ARTICLE

On preferences with infinitely many subjective states

Kalyan Chatterjee · R. Vijay Krishna

Received: 7 May 2009 / Accepted: 22 September 2009

Abstract Models with subjective state spaces have been extremely useful in capturing novel psychological phenomena that consist of both a preference for flexibility and for commitment. Interpreting the utility representations of preferences as capturing these phenomena requires one to use the notion of a sign of a state. For linear preferences, we completely characterise the sign of a state in terms of its analytic representation as an integral with respect to a signed measure. In models with finitely many states, a state is either positive or negative, but never both. We show that in models with infinitely many states, a state can be both positive and negative. Thus, models with finitely many states may not capture all the behavioural features of an infinite model. Our methods are also useful in constructing utility functionals over menus with desired local properties.

Keywords Subjective state space · Temptation · Preference for flexibility · Preference for commitment

JEL Classification D81

We would like to thank an anonymous referee for useful comments that greatly improved the exposition.

K. Chatterjee
Department of Economics, The Pennsylvania State University,
609 Kern Building, University Park, PA 16802-3306, USA
e-mail: kchatterjee@psu.edu

R. V. Krishna (✉)
Department of Economics CB 3305, University of North Carolina,
Chapel Hill, NC 27599, USA
e-mail: rvk@email.unc.edu

1 Introduction

Since the seminal paper of Dekel et al. (2001) (henceforth DLR), there has been a great deal of interest in the subjective states that an economic agent considers possible. Roughly put, by looking at an agent's preferences over planning problems, we can make inferences about how she (the agent) thinks she will behave in the future. The set of future 'states of mind' that the agent considers possible constitute her *subjective state space*, with the understanding that two states of mind are different if (and only if) they correspond to her behaving differently in the two states.

A *subjective state* therefore corresponds to a utility function according to which the agent will make a choice in the future.[1] The identification of the subjective state space has been used to capture novel psychological phenomena such as temptation (Gul and Pesendorfer 2001; Dekel et al. 2008), non-Bayesian updating (Epstein 2006), regret (Sarver 2008) and perfectionism and time-varying pessimism (described in Kopylov 2008). All the papers mentioned above, except for Sarver (2008), assume (or make assumptions so) that the subjective state space of the agent is finite.

DLR consider planning problems that are represented by menus (i.e. closed sets) of lotteries over a set of prizes. They show that if preferences over planning problems (i.e. menus of lotteries) are linear, there exists a representation of the preference as the integral of a finite, signed measure on the subjective state space. Such a representation is referred to as an *additive EU representation* and is described in Sect. 2.

A useful notion, also introduced by DLR and put to use in the literature, is the notion of the sign of a state. We may attach a sign to a subjective state, wherein, a positive state indicates a (local) preference for flexibility, i.e. more options are better, and a negative state indicates a preference for commitment, i.e. fewer options are better. Negative states lie at the heart of phenomena such as temptation and regret. (More formal definitions are provided in Sect. 2.) While it is straightforward to *define* the sign of a state, one may ask

Given an additive EU representation, what is the sign of any specific state?

In Theorem 6 we provide a precise characterisation of the sign of a state in terms of the signed measure representing the preference. In particular, we provide a constructive answer to the question above. An interesting consequence of our theorem is that even if preferences over menus are linear, a state can be both positive and negative.

If the state space is finite (and this is a property of the preference), every state is either positive or negative, but never both. This is a very useful property since it allows us to unambiguously interpret the various negative states as representing, for instance, temptation or regret, while the positive states represent a preference for flexibility.

We shall, however, show in this paper that if the subjective state space is infinite (possibly countable), then a state can simultaneously be both positive and negative, even if preferences are linear over planning problems.[2]

---

[1] Thus, the subjective state space is a subset of the space of all twice normalised vN-M functions on Z.

[2] DLR claim (p. 912) that if preferences are linear, then regardless of the cardinality of the subjective state space, a state is either positive or negative, but not both. Our example in Sect. 2.4 demonstrates this is not true.

This result suggests that the sign of a state is perhaps not the most useful concept when the state space is infinite. Apart from Sarver (2008) who allows for an infinite subjective state space and Gul and Pesendorfer (2001) who make behavioural assumptions so that the subjective state space is finite, most of the other papers that use a finite state space do so for analytical convenience. Indeed, if the state space is infinite, analytic results are typically hard to obtain; hence the assumption that the state space is finite.

Our results can also be interpreted as making a case for infinite subjective state spaces, unless natural behavioural restrictions preclude them, since while it is possible to approximate preferences with an infinite subjective state space with a sequence of preferences with finitely many subjective states in utility terms, the approximations may not capture all the behavioural properties of the original preference, thus highlighting the subtleties that lie in infinite state spaces. From a pedagogical viewpoint, our contribution is in showing how to construct preferences over menus that have certain desirable local properties, even when the subjective state space is infinite. This is useful in constructing examples and, as is the case here, counterexamples.

We shall proceed as follows: In Sect. 2, we introduce the model, in Sect. 2.1 we introduce two classes of continuous functions that will prove useful to us, in Sect. 2.3 we describe the intuition behind our results, in Sect. 2.4 we present the aforementioned example and in Sect. 2.5, we present a theorem which classifies a state as being positive or negative in terms of the utility representation. We discuss some alternative definitions of the sign of a state in Sect. 2.6 and conclude in Sect. 3. We now provide a brief review of the literature.

Related literature

The paper by DLR has been extremely influential as it gives conditions for the existence of a unique subjective state space, without requiring preferences to be monotone. An alternate approach to the subjective state space is in Sagi (2006). Non-monotone preferences have been the subject of many of the recent investigations. Gul and Pesendorfer (2001) were the first to show that the environment of menus of lotteries is ideal for capturing phenomena such as temptation and self-control. Since then, numerous papers such as Dekel et al. (2008) and Epstein (2006) have used the same basic model to capture various psychological phenomena. Kopylov (2008) provides a unified axiomatic treatment of subjective state space models with finitely many states. His main axiom (which ensures finiteness) is novel and very useful, as it applies in more general, nonlinear contexts as well. All of the papers mentioned above assume that the subjective state space is finite, and this assumption is made for the analytical convenience it affords.

Sarver (2008) considers the effect that the anticipation of (ex-post) regret has on preferences over menus (planning problems). His work allows for infinitely many subjective states. He also provides a very careful comparison of his model, where he studies the consequences of anticipating regret, with the extant literature. We should point out that in Sarver's paper, there is only one positive state. In the finite state version of his model, this positive state can never be negative, but there are instances of his model with infinitely many states, where the positive state can also be negative. (Sarver's representation consists of a probability measure over states and a Dirac point measure at a point (of opposite sign). Our example in Sect. 2.4 captures essentially this feature.) We discuss his model in greater length in Sect. 2.2.

2 The model

The description of the model follows DLR closely, except for some notational differences. Let $Z := \{1, \ldots, n+1\}$ be a set of prizes and $\Delta^n := \left\{x \in \mathbb{R}_+^{n+1} : \sum_{i=1}^{n+1} x_i = 1\right\}$ the space of lotteries over Z. Let d be the Euclidean metric on Δ^n, thereby generating the Euclidean topology. The space of all compact subsets of Δ^n is denoted by $\mathcal{F}(\Delta^n)$. For any $A, B \in \mathcal{F}(\Delta^n)$, the Hausdorff distance between A and B is

$$\mathrm{d}_h(A, B) = \max\left\{\max_{x \in A} \min_{y \in B} \mathrm{d}(x, y), \max_{y \in B} \min_{x \in A} \mathrm{d}(x, y)\right\}.$$

When endowed with the Hausdorff metric, $\mathcal{F}(\Delta^n)$ becomes a compact metric space. A *menu* is a closed subset of Δ^n. A preference $\succcurlyeq$ is a complete, transitive binary relation on $\mathcal{F}(\Delta^n)$. A preference $\succcurlyeq$ satisfies *Independence* if $A \succ B$ implies $\lambda A + (1-\lambda)C \succ \lambda B + (1-\lambda)C$ for all $\lambda \in (0, 1]$ and for all menus A, B, C. A preference satisfies *Indifference to Randomisation (IR)* if for every menu A, $A \sim \mathrm{conv}\,(A)$ (where $\mathrm{conv}\,(A)$ is the convex hull of A). In what follows, we shall assume that preferences are continuous in the Hausdorff metric, i.e. for each $A \in \mathcal{F}(\Delta^n)$, the upper and lower contour sets, namely $\{B : B \succcurlyeq A\}$ and $\{B : A \succcurlyeq B\}$ are closed. Lemma 1 of DLR states that if a preference is continuous and satisfies Independence, then it also satisfies IR. Therefore, in what follows, we shall restrict attention to compact convex subsets of Δ^n. The space of all compact, convex subsets of Δ^n is denoted by $\mathcal{K}(\Delta^n)$.

There is an implicit time line here. In the morning, the agent makes a choice *of* a menu, knowing that in the evening, she will make a choice *from* the menu. The agent also knows that her choice in the evening will be made according to a vN-M utility function. Thus, the agent has preferences over planning problems, and her preference over menus enables us to infer how she believes she will feel, i.e. what subjective state of mind she thinks she will be in the evening, when she makes her choice from the menu.

Let S_Z be the space of vN-M utility functions on Δ^n. With two normalisations, $S_Z := \left\{s \in \mathbb{R}^{n+1} : \sum_i s_i = 0 \text{ and } \sum_i s_i^2 = 1\right\}$ can be identified with S^{n-1}, the $(n-1)$-dimensional sphere. S_Z constitutes the set of all possible subjective states that the agent can be in when making a choice from the menu. The utility in state s is given by $u(\cdot, s)$, and the utility from the lottery $x \in \Delta^n$ is $u(x, s) = \langle x, s\rangle$ (where $\langle \cdot, \cdot \rangle$ is the standard inner product in $\mathbb{R}^{n+1}$).

Following DLR (p. 909), we adopt the following definitions. A state $s \in S_Z$ is *positive* if for each neighbourhood N of s, there exist menus A and B such that $A \subset B$ and $B \succ A$, wherein $\max_{x \in A} u(x, s') = \max_{x \in B} u(x, s')$ for all $s' \in S_Z \setminus N$. The intuition behind this definition is simple. Fix a menu A and a state s. Then, if increasing the ex-post utility for some utility functions in a neighbourhood of s (thereby obtaining menu B) makes the agent better off, the state is positive. Similarly, a state $s \in S_Z$

is *negative* if for each neighbourhood N of s, there exist menus A and B such that $A \subset B$ and $A \succ B$, wherein $\max_{x \in A} u(x, s') = \max_{x \in B} u(x, s')$ for all $s' \in S_Z \setminus N$.

Roughly put, positive states indicate that the agent (locally) prefers flexibility and negative states indicate that he prefers fewer choices, i.e. commitment. Thus, the sign of a state has behavioural significance, as it measures how a local change to a menu impacts the agent's utility. (We emphasise that the definition of a sign of a state does *not* depend on any assumptions on preferences, beyond continuity and IR. Continuous preferences, that satisfy IR possess *weak EU representations*, DLR, p. 903.)

We now recall some definitions, which may also be found in Dudley (2003). The Borel σ-algebra on S_Z is the σ-algebra generated by the open subsets of S_Z, and is denoted by $\mathcal{B}(S_Z)$. A finite regular Borel measure is a set function $\mu : \mathcal{B}(S_Z) \to \mathbb{R}_+$ defined on the Borel σ-algebra that is non-negative, countably additive, finite (in the sense that $\mu(S_Z) < \infty$) and satisfies $\mu(A) := \inf \{U : A \subset U \in \mathcal{B}(S_Z),\ U \text{ open}\} = \sup \{C : A \supset C \in \mathcal{B}(S_Z),\ C \text{ closed}\}$, i.e. the measure of a set A can be approximated from the outside by open sets that contain A, or by closed sets contained in A.

A *signed* finite regular Borel measure $\mu : \mathcal{B}(S_Z) \to \mathbb{R}$ is a set function that can also take negative values. Every signed measure μ has a Hahn–Jordan decomposition $\mu := \mu^+ - \mu^-$, where μ^+ and μ^- are (positive) measures such that $\mu^+ \perp \mu^-$, i.e. μ^+ and μ^- are mutually singular. In other words, μ^+ and μ^- are positive on disjoint subsets of the domain.

For a (positive) measure μ, we may define its support as $\text{supp}(\mu) := \text{cl}\{s \in S_Z : s \in U, U \text{ open, implies } \mu(U) > 0\}$ (where $\text{cl}\,(A)$ indicated the closure of the set A). Therefore, for any signed measure $\mu = \mu^+ - \mu^-$, the support of μ is $\text{supp}(\mu) := \text{supp}(\mu^+) \cup \text{supp}(\mu^-)$.

Dekel et al. (2001, 2007) show that a (Lipschitz) continuous preference that satisfies Independence can be represented by a linear utility function $V : \mathcal{K}(\Delta^n) \to \mathbb{R}$ that has the form

$$V(A) := \int_{S_Z} \max_{x \in A} u(x, s)\ \mu(\mathrm{d}s),$$

where μ is a finite, signed, regular Borel measure on S_Z (with the Borel σ-algebra) and $u(\cdot, s) \in S_Z$ is a vN-M utility function on Δ^n for each $s \in S_Z$. Such a representation is referred to as an *additive EU representation*. In what follows, we shall refer to preferences and the corresponding measure μ interchangeably.

We shall refer to the support of the measure μ as the *subjective state space*. Notice that this may be infinite and, in general, its cardinality depends on the preference in question. Consider now a preference represented by a utility function $V(A) = \int_{S_Z} \max_{x \in A} u(x, s)\ \mu(\mathrm{d}s)$. Suppose also that μ has infinite support. By standard arguments, it follows that there exists a sequence of measures with finite support (μ_n) such that the corresponding (linear) utility functions V_n converge to V.

Consider now an additive EU representation with a finite state space. Since the measure μ is a signed measure, it has a Hahn–Jordan decomposition $\mu := \mu^+ - \mu^-$, where μ^+ and μ^- are (positive) measures such that $\mu^+ \perp \mu^-$. If the state space is finite, then $P := \operatorname{supp}(\mu^+)$ and $N := \operatorname{supp}(\mu^-)$ are finite. Then, the utility representation can be written as

$$V(A) := \sum_{s \in P} \alpha_s \max_{x \in A} u(p, s) - \sum_{s \in N} \alpha_s \max_{x \in A} u(p, s),$$

where $\alpha_s := |\mu(s)| > 0$ for all $s \in S_Z$. The set P represents the positive states, and S the negative states. Clearly, no state is both positive and negative.

In the following sections, we shall construct a preference $\succcurlyeq$ with a subjective state that is both positive and negative. We begin with some preliminaries about two useful classes of continuous functions.

2.1 Support and bump functions

We will now describe two very useful classes of functions, support and bump functions. Support functions provide a dual characterisation of convex sets, and spaces of convex sets then correspond to spaces of support functions. This proves to be analytically far more convenient than dealing with the convex sets themselves. Bump functions will prove extremely useful in constructing the desired menus, i.e. compact, convex subsets of Δ^n. Roughly speaking, a bump function is a smooth approximation of the indicator function of a compact set. Now for some specifics.

Associated with a compact, convex subset A of Δ^n is its *support function* $h_A : S_Z \to \mathbb{R}$. Support functions are useful since there exists a natural isomorphism between the space of compact, convex subsets of Δ^n and the space of support functions. Our discussion of support functions follows DLR, but our notation is a little different.

As above, we let $S_Z := \left\{ s \in \mathbb{R}^{n+1} : \sum_i s_i = 0 \text{ and } \sum_i s_i^2 = 1 \right\}$ be the set of normalised set of vN-M utility functions on Δ^n. For $A \in \mathcal{K}(\Delta^n)$, the support function $h_A : S_Z \to \mathbb{R}$ is defined as $h_A(s) := \max_{x \in A} \langle x, s \rangle$. (Notice that with this definition, $h_A(s) = \max_{x \in A} u(x, s)$.) Thus, for any additive EU functional $V : \mathcal{K}(\Delta^n) \to \mathbb{R}$ and menu A, $V(A) = \int_{S_Z} h_A(s)\, \mu(\mathrm{d}s)$.) For more details about support functions, see Schneider (1993).

Intuitively, each $s \in S_Z$ provides a *direction*, and $h_A(s)$ is the largest value taken on A by a hyperplane that has s normal to it, so the hyperplane increases in the direction of s, and all its level sets are orthogonal to s. This also suggests that knowing the value of the support function in every direction completely describes the set A (up to closure). This is indeed the case, as we shall see below.

As always, the Banach space of continuous functions on S_Z is $C(S_Z)$, with the supremum norm, whereby $\|f\|_\infty = \max_{s \in S_Z} |f(s)|$ for all $f \in C(S_Z)$. We may define an order on $C(S_Z)$ using the pointwise order: for $f, g \in C(S_Z)$, $f \geq g$ if $f(s) \geq g(s)$ for all $s \in S_Z$. Let $K^* := \{h_A \in C(S_Z) : A \in \mathcal{K}(\Delta^n)\}$, and for any $h \in K^*$, let

$$A_h := \bigcap_{s \in S_Z} \left\{ x \in \Delta^n : \langle x, s \rangle \leq h(s) \right\}.$$

Support functions have the following useful properties. (i) $h_{\lambda A+(1-\lambda)B} = \lambda h_A + (1-\lambda)h_B$, (ii) $h_A \wedge h_B = h_{A\cap B}$, (iii) $h_A \vee h_B = h_{\text{conv}\,(A\cup B)}$, (iv) $A \subset B$ if and only if $h_A \leq h_B$ and (v) $\|h_A - h_B\|_\infty = d_h(A, B)$. Also, the following duality relation holds: $h_{A_h} = h$ and $A_{h_A} = A$. Thus, the support function is a complete description of a compact, convex set.

The convexity of the set K^* follows from the above properties. To see that $\mathbf{0} \in K^*$, recall the definition of S_Z, wherein $\sum_i s_i = 0$ for all $s \in S_Z$. Then, $\langle s, (\frac{1}{n+1}, \ldots, \frac{1}{n+1})\rangle = 0$, which implies that $h_{\left(\frac{1}{n+1},\ldots,\frac{1}{n+1}\right)}(s) = 0$ for all $s \in S_Z$. Thus, $\mathbf{0} = h_{\left(\frac{1}{n+1},\ldots,\frac{1}{n+1}\right)} \in K^*$. In summary, $K^* \subset C(S_Z)$ is a closed, convex subset of $C(S_Z)$ that contains the origin $\mathbf{0}$.

Let $\text{cone}(K^*) := \bigcup_{\lambda\geq 0} \lambda K^*$ be the convex cone generated by K^*. Intuitively, $\text{cone}(K^*)$ corresponds to the space of compact, convex subsets of $\text{aff}(\Delta^n)$ (where $\text{aff}(\Delta^n)$ is the affine hull of Δ^n). In particular, for any $f \in \text{cone}(K^*) \subset C(S_Z)$, define

$$\mathrm{K_f} := \bigcap_{s\in S_Z} \left\{x \in \text{aff}(\Delta^n) : \langle x, s\rangle \leq \mathrm{f(s)}\right\}.$$

Suppose $f \in \text{cone}(K^*) \setminus K^*$. Then, there exists $h \in K^*$ and $\lambda > 0$ such that $h = \lambda f = \lambda f + (1-\lambda)\mathbf{0}$. Moreover, $K_h = K_{\lambda f+(1-\lambda)\mathbf{0}} = \lambda K_f + (1-\lambda)\left(\frac{1}{n+1}, \ldots, \frac{1}{n+1}\right)$. Similarly, let $K \in \text{aff}\Delta^n$. Then, there exists $\lambda \in (0, 1)$ such that $C := \lambda K + (1-\lambda)\left(\frac{1}{n+1}, \ldots, \frac{1}{n+1}\right) \subset \Delta^n$. Thus, $h_C = h_{\lambda K+(1-\lambda)\left(\frac{1}{n+1},\ldots,\frac{1}{n+1}\right)}$, i.e. $h_K = (1/\lambda)h_C$. Thus, $f \in \text{cone}(K^*)$ if and only if $K_f \in \text{aff}\Delta^n$.

Thus (see also Schneider 1993), there is an isometry between the compact, convex subsets of $\text{aff}(\Delta^n)$ and $\text{cone}(K^*)$. Moreover, a function $f \in C(S_Z)$ is a support function of some compact, convex subset of $\text{aff}(\Delta^n)$ if and only if $f^* : \mathbb{R}^n \to \mathbb{R}$, the unique extension of f to $\mathbb{R}^n$ ($\simeq \text{span}(S_Z)$) by positive homogeneity, is sublinear. The following lemma, which is Lemma 1.7.9 in Schneider (1993), will be extremely useful.

Lemma 1 *Let $f \in C(S_Z)$ be twice continuously differentiable. Then, there exists $r > 0$ such that $f + r\mathbf{1}$ is the support function of some compact, convex set in $\mathbb{R}^n$.*

Recall that twice differentiable functions are dense in $C(S_Z)$, in the sense that for any $f \in C(S_Z)$ and for any $\varepsilon > 0$, there exists a twice differentiable function $g \in C^2(S_Z)$ such that $\|f - g\|_\infty < \varepsilon$. Thus, $\text{span}(K^*)$ is dense in $C(S_Z)$.

We now turn to another class of functions that will be useful for our purposes, the so-called *bump functions*. As mentioned above, bump functions allow us to smoothly approximate indicator functions of sets. We begin with the smooth function

$$\alpha(t) := \begin{cases} \mathrm{e}^{-1/t} & \text{if } t > 0, \\ 0 & \text{if } t \leq 0. \end{cases}$$

Now define the function

$$g(t; \varepsilon) := \frac{\alpha(t)}{\alpha(t) + \alpha(\varepsilon - t)},$$

where $\varepsilon > 0$. Observe that for any $t \in \mathbb{R}$, the denominator is strictly positive and so g is smooth. Moreover, for $t \leq 0$, $g(t) = 0$, for $t > 0$, $g(t) > 0$ and for $t \geq \varepsilon$, $g(t) = 1$. Also, g is non-decreasing.

We are now ready to define our first bump function $\varphi(\cdot, \varepsilon) : \mathbb{R} \to \mathbb{R}$ as $\varphi(t; \varepsilon) := g(\varepsilon - |t|)$. The following are important properties of $\varphi(\cdot; \varepsilon)$. (i) For $t \notin [-\varepsilon, \varepsilon]$, $\varphi(t; \varepsilon) = 0$, $\varphi(0; \varepsilon) = 1$, (ii) $\varphi(\cdot; \varepsilon)$ is smooth, and (iii) for $t \in [-\varepsilon, \varepsilon]$, $\varphi(t; \varepsilon) = \varphi(-t; \varepsilon)$, so $\varphi(\cdot; \varepsilon)$ is symmetric about 0. By translation, we can define a bump function at any point.

Another bump function useful for us is $\psi(\cdot; x, \varepsilon) : \mathbb{R} \to \mathbb{R}$, defined as $\psi(t; x, \varepsilon) := g(-|t| + x; \varepsilon)$ where $0 < \varepsilon < x$. Notice that $\psi(t; x, \varepsilon) = \psi(-t; x, \varepsilon)$ for all $t \in \mathbb{R}$, $\psi(t; x, \varepsilon) = 0$ for $|t| \geq x$, $\psi(t; x, \varepsilon) = 1$ for $|t| < x - \varepsilon$ and $\psi(\cdot; x, \varepsilon)$ is smooth. More generally, the following is true.

Proposition 2 *Let M be a smooth (finite dimensional) manifold (without boundary) and $K \subset U \subset M$ where K is compact and U is open. Then, there exists a smooth bump function $\varphi_M : M \to \mathbb{R}$ denoted by $\varphi_M(x; U, K)$ such that $\varphi_M(x; U, K) = 0$ for all $x \in U^c$ and $\varphi_M(x; U, K) = 1$ for all $x \in K$.*

We should note that as defined, S_Z is a smooth $(n - 1)$-dimensional manifold without boundary. The advantage of this proposition is that U can be taken to be arbitrarily close to K in any suitable sense. Conversely, given an open set U, one may choose a compact set $K \subset U$ where K and U can be taken to be arbitrarily close in a suitable sense. The proposition is Proposition 2.26 (p. 55) of Lee (2003), where a proof may also be found.

2.2 Regret and commitment preference

Sarver studies an agent who must make a choice from a menu and anticipates ex-post regret. He introduces a *regret representation* which is a pair (μ, K) where $K \geq 0$ and μ is a probability measure on S_Z. (Our notation differs slightly from his.) The utility of a menu is then

$$V(A) = \max_{p \in A} \int\limits_{S_Z} [u(p, s) - R(p, A, s)] \, \mathrm{d}\mu(s),$$

where

$$R(p, A, s) = K \left[\max_{q \in A} u(q, s) - u(p, s) \right],$$

where $R(p, A, s)$ is the ex-post regret from choosing p from the menu A, when subjective state s has been realised. The interpretation of this representation is that the agent has subjective uncertainty about her tastes, but must make a choice (from a menu) before her subjective state is realised. She acts according to V, trading off expected utility and ex-post regret. It is easily seen that the representation above can be rewritten as

$$V(A) = \max_{p \in A} \left[(1+K) \int_{S_Z} u(p,s)\, \mathrm{d}\mu(s) \right] - K \int_{S_Z} \max_{q \in A} u(q,s)\, \mathrm{d}\mu(s).$$

Notice that this is not the interpretation of the DLR model. Nevertheless, such a V also has an additive EU representation. This is Lemma 1 of Sarver (2008), where it is noted that $(1+K) \int_{S_Z} u(\cdot, s)\, \mathrm{d}\mu(s)$ is an expected utility function, so that there exists $\alpha \geq 0$ and $s^* \in S_Z$ such that $\alpha u(\cdot, s^*) = (1+K) \int_{S_Z} u(\cdot, s)\, \mathrm{d}\mu(s)$. Thus, the value of a menu is

$$V(A) = \alpha \max_{p \in A} u(p, s^*) - K \int_{S_Z} \max_{q \in A} u(q,s)\, \mathrm{d}\mu(s),$$

which is easily seen to be an additive EU representation.

Intuitively, s^* is a positive state and represents the agent's *commitment preference*. Moreover, every $s \in \mathrm{supp}(\mu) \setminus \{s^*\}$ is a negative state. Roughly put, this is because the states in $\mathrm{supp}(\mu) \setminus \{s^*\}$ represent deviations from the agent's commitment choice. As Sarver (2008, p. 272) notes, negative states are '… key to capturing regret.' We shall show that if $s^* \in \mathrm{supp}(\mu)$, s^* too is negative. Thus, s^* is both positive and negative, which complicates its interpretation as the commitment preference. As it turns out, the regret in state s^* is 0, but the state is nonetheless negative.

Example 3 Suppose $Z := \{1, 2, 3\}$. Then, $S_Z \simeq S^1 \simeq [0,1]/\{0,1\}$ (where we identify $\{0\}$ and $\{1\}$, and use the quotient topology). Fix $K > 0$, $\delta < \pi/2$ and μ uniform on $\left[\frac{\pi}{2} - \delta, \frac{\pi}{2} + \delta\right]$. Then, $s^* = \frac{\pi}{2}$. In Sect. 2.4, we demonstrate that s^* is both positive and negative. This also follows from the theorem in Sect. 2.5.

2.3 An overview of our results

We shall now take stock of all the ideas introduced thus far, and briefly sketch how they will be used. Recall the definition of a positive state. A state $s^* \in S_Z$ is *positive* if for each neighbourhood N of s^*, there exist menus A and B such that $A \subset B$ and $B \succ A$, wherein $\max_{x \in A} u(x,s) = \max_{x \in B} u(x,s)$ for all $s \in S_Z \setminus N$. But $\max_{x \in A} u(x,s) = h_A(s)$ for all $s \in S_Z \setminus N$, so it is required that $h_A(s) = h_B(s)$ for all $s \in S_Z \setminus N$. To show that s^* is positive, we want to find menus A and B whose support functions differ *only* on N, such that $\int_{S_Z} [h_A(s) - h_B(s)]\, \mathrm{d}\mu(s) > 0$.

We will show that a state s^* is positive if, and only if, $s^* \in \mathrm{supp}(\mu^+)$ (with a similar result for negative states). Suppose $s^* \in \mathrm{supp}(\mu^+)$. Then, there exists an open neighbourhood $U \ni s^*$ such that $U \cap \mathrm{supp}(\mu^+) \neq \varnothing$. (We have used here the fact that $\mu^+ \perp \mu^-$.) Therefore, there exists a bump function $f : S_Z \to \mathbb{R}$ such that $f = 0$ outside of U, but $\int_{S_Z} f(s)\, \mathrm{d}\mu(s) = \int_U f(s)\, \mathrm{d}\mu(s) > 0$. (That such a U and f exist is established in Theorem 6.) If f is the difference of two support functions, then we are done, but this is frequently not the case.

This is where we use Lemma 1, which says that there exists an $r > 0$ such that $f + r\mathbf{1}$ is the support function of some convex body. Let $K_{f+r\mathbf{1}}$ be this body. Notice that this convex body contains the body $K_{r\mathbf{1}}$, which is the ball of radius r, around $\left(\frac{1}{n+1}, \dots, \frac{1}{n+1}\right)$. Now, let $B = K_{f+r\mathbf{1}}$, $A = K_{r\mathbf{1}}$, and observe that $A \subset B$ and $V(B) - V(A) = \int_{S_Z} f(s)\,\mathrm{d}\mu(s) > 0$, which is equivalent to $B \succ A$. The case where $s^* \in \operatorname{supp}(\mu^-)$ is handled similarly.

It is worthwhile to mention the key difference between finite and infinite subjective state spaces. If the subjective state space is infinite, and the subjective state s^* is not isolated (in the sense that every neighbourhood of s^* contains another subjective state other than s^*), then the set U that we have chosen, where $f > 0$, need not even contain s^*. Clearly, in the finite case, this cannot not be so, i.e. U must contain s^* (since every state must be isolated). This is the essence of our observations.

Put differently, we show three things: First, we show that a state s^* is positive if, and only if, $s^* \in \operatorname{supp}(\mu^+)$; second we show that a state s^* is negative if, and only if, $s^* \in \operatorname{supp}(\mu^-)$, and finally, we show that it is possible that $\operatorname{supp}(\mu^+) \cap \operatorname{supp}(\mu^-) \neq \varnothing$. In Sect. 2.4, we go through the mechanics in a simple example with three prizes. In Sect. 2.5 we fill in the details for arbitrary (but finite) Z.

2.4 An example

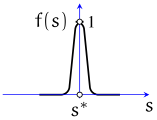
We now provide an example of a preference that admits an additive EU representation, but has a subjective state that is both positive and negative. We shall define the preference $\succcurlyeq$ by constructing a function $V : K^* \to \mathbb{R}$ and recalling the isomorphism between K^* and $\mathcal{K}(\Delta^n)$. More precisely, we shall define a linear functional $V : C(S_Z) \to \mathbb{R}$ so that restricting the function to $K^* \subset C(S_Z)$ provides the desired function.

For notational simplicity only, we shall consider an example where $n = 2$, so that $S_Z \simeq S^1 \simeq [0, 1]/\{0, 1\}$. A general example for arbitrary n will be seen to follow immediately, but at the cost of additional notation. Moreover, our characterisation theorem, Theorem 6, handles the general case. The example below merely provides the intuition.

Fix $s^* \in [0, 1) = S_Z$. Let μ be a signed measure on S_Z with Jordan decomposition $\mu := \mu^+ - \mu^-$, where μ^+ is the uniform measure on S_Z and μ^- is a Dirac measure concentrated at s^*. (This means μ^- is defined as follows: $\mu^-(A) = 1$ if $s^* \in A$, $\mu^-(A) = 0$ if $s^* \notin A$.) Define the linear functional $V : C(S_Z) \to \mathbb{R}$ as $V(f) := \int_{S_Z} f(s)\,\mu(\mathrm{d}s)$ for all $f \in C(S_Z)$. Then, for menus $A, B \in \mathcal{K}(\Delta^2)$, $A \succcurlyeq B$ if (and only if) $V(h_A) \geq V(h_B)$, where $h_A, h_B \in K^*$.

Claim 4 The state s^* is negative.

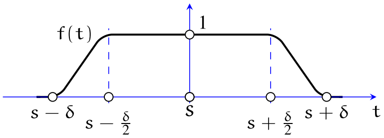
Proof Let $\delta > 0$ such that $\delta < \mu^-(s^*)/2 = 1/2$ and $[s^* - \delta, s^* + \delta] \subset [0, 1)$, and define $f : [0, 1) \to \mathbb{R}$ as follows. For $s \in [s^* - \delta, s^* + \delta]$, let $f(s) := \varphi(s; \delta)$, so f is the bump function defined above centred at s^* (see picture below).



Recall that $\mu^+([s^* - \delta, s^* + \delta]) = 2\delta$. Also, $f(s) = 0$ for $s \notin [s^* - \delta, s^* + \delta]$. Notice that $V(f) = \int f(s)\, \mathrm{d}\mu^+(s) - \int f(s)\, \mathrm{d}\mu^-(s) < 2\delta - 1\mu^-(s^*) < 0$. Notice that $f \geq \mathbf{0}$, so that for all $r > 0$, $f + r\mathbf{1} \geq \mathbf{0} + r\mathbf{1}$. Now, by Lemma 1, there exists $r > 0$ such that $f + r\mathbf{1}$ is the support function of some compact convex body. Thus, the convex body $K_{f+r\mathbf{1}} \supset K_{r\mathbf{1}} (= B_r)$ where B_r is the ball of radius r. Also, $V(f + r\mathbf{1}) = V(f) + V(r\mathbf{1}) < V(r\mathbf{1})$. Notice that $K_{f+r\mathbf{1}}$ need not be contained in Δ^n, but there exists $\lambda > 0$ such that $K_{\lambda(f+r\mathbf{1})} = \lambda K_{f+r\mathbf{1}} + (1-\lambda)\left(\frac{1}{n+1}, \ldots, \frac{1}{n+1}\right) \subset \Delta^n$. Then, $V(\lambda(f + r\mathbf{1})) < V(\lambda r\mathbf{1})$, so that $B_{\lambda r} \succ \lambda K_{f+r\mathbf{1}}$. Since the construction works for any neighbourhood of s^*, we see that s^* is negative. □

Claim 5 The state s^* is positive.

Proof Let N be a neighbourhood of s^* and $\varepsilon > 0$ such that $[s^* - \varepsilon, s^* + \varepsilon] \subset N$. Fix some $s \in [s^* - \varepsilon, s^* + \varepsilon]$ and define $f : [0, 1) \to \mathbb{R}$ as follows. For $t \in [s - \delta, s + \delta]$ where $\delta < |s - s^*|$, let $f(t) := \psi(t - s; \delta, \delta/2)$. Therefore, $f(t) = 1$ for $t \in [s - \delta/2, s + \delta/2]$, $f(t) = 0$ for $t \notin [s - \delta, s + \delta]$, $f \geq \mathbf{0}$ and f is smooth. Moreover, $f(s^*) = 0$.



Since $f \geq \mathbf{0}$, by Lemma 1, there exists $r > 0$ such that $f + r\mathbf{1}$ is the support function of a compact, convex body, and $f + r\mathbf{1} \geq r\mathbf{1}$. Again, $V(f) = \int_{S_Z} f(t)\, \mu(\mathrm{d}t) > \int_{s-\delta/2}^{s+\delta/2} f(t)\, \mu^+(\mathrm{d}t) = \delta > 0$, so that $V(f + r\mathbf{1}) = V(f) + V(r\mathbf{1}) > V(r\mathbf{1})$. Arguing as above, we see that the state s^* is positive. □

2.5 A theorem

We now present a theorem which tells us the sign of a state given the signed measure that represents the preference. The set of prizes is Z, a finite set. Recall that a state $s \in S_Z$ (defined above) is *positive* if for each neighbourhood N of s, there exist menus A and B such that $A \subset B$ and $B \succ A$, wherein $\max_{x \in A} u(x, s') = \max_{x \in B} u(x, s')$ for all $s' \in S_Z \setminus N$, with a similar definition for negative states. Let $\mathcal{P} := \mathrm{cl}\,\{s \in S_Z : s \text{ is positive}\}$ be the closure of the set of positive states and $\mathcal{N} := \mathrm{cl}\,\{s \in S_Z : s \text{ is negative}\}$ be the closure of the set of negative states (where cl A indicates the closure of the set A). The definitions of $\mathcal{P}$ and $\mathcal{N}$ follow DLR (p. 909).

As before, μ is a signed measure on S_Z and $\mu := \mu^+ - \mu^-$ is its Hahn–Jordan decomposition. Let $C_+ := \text{supp}(\mu^+)$ and $C_- := \text{supp}(\mu^-)$ be the supports of μ^+ and μ^-, respectively. Recall that by definition of the support of a measure, C_+ and C_- are closed. Again, $V : C(S_Z) \to \mathbb{R}$ is given by $V(f) := \int f(t)\, \mu(\mathrm{d}t)$. For menus A and B, $A \succcurlyeq B$ if and only if $V(h_A) \geq V(h_B)$.

Theorem 6 *A state s is positive if and only if $s \in C_+$, and a state s is negative if and only if $s \in C_-$. In other words, $\mathcal{P} = C_+$ and $\mathcal{N} = C_-$.*

Proof We shall first show that if $s \in C_+ \setminus C_-$, then s is positive but not negative. We shall then show that if $s \in C_+ \cap C_-$, then s is positive. Replacing positive with negative and changing some signs and the direction of some inequalities completes the proof of the theorem.

Suppose $s \in C_+ \setminus C_-$. Then, there exists an open set $U_s \ni s$ such that $U_s \cap C_- = \varnothing$. Furthermore, for any open $U \subset U_s$, it is still the case that $U \cap C_- = \varnothing$. Fix any open $U \subset U_s$ that contains s.

To see that s is not negative, let $A, B \in \mathcal{K}(\Delta^n)$ such that $A \subset B$, and so that $h_A(t) = h_B(t)$ for $t \in U_s^c$. But $h_A \geq h_B$, so that $\int (h_A - h_B)\, \mu(\mathrm{d}t) \geq 0$. Thus, s is not negative.

To show that s is positive, let $K \subset U$ be compact such that $K \ni s$. Then, by Proposition 2 above, there exists a bump function $\psi_{S_Z}(t; U, K)$ such that ψ_{S_Z} is 0 on U^c and 1 on K. Let $f(t) = \psi_{S_Z}(t; U, K)$, so that $V(f) = \int f(t)\, \mu(\mathrm{d}t) > 0$ (since $U \cap C_- = \varnothing$). Since f is smooth, there exists $r > 0$ such that $f + r\mathbf{1}$ is the support function of a convex body. Then, $V(f + r\mathbf{1}) = V(f) + V(r\mathbf{1}) > V(r\mathbf{1})$. Following the arguments above in the example, we see that s is positive.

We now show that if $s \in C_+ \cap C_-$, then s is positive. We first prove a useful bounding result on the Hahn–Jordan decomposition. Let $C_0 := C_+ \cap C_-$, and fix $U_0 \ni s$ open, so that $U_0 \cap C_0 \neq \varnothing$. Let $\mathcal{I} := \{U \subset U_0 : U \ni s,\ U \text{ open}\}$. We claim that $\sup_{U \in \mathcal{I}} \mu(U) > 0$ and $\inf_{U \in \mathcal{I}} \mu(U) < 0$.

We shall only prove the first part of the claim. To see that the claim is true, suppose not, so that for all $U \in \mathcal{I}$, $\mu(U) \leq 0$. Then, $\mu^-(U) \geq \mu^+(U)$ for all $U \in \mathcal{I}$. But this is impossible, given the Hahn–Jordan decomposition of μ, whereby $\mu^+ \perp \mu^-$ and the fact that μ is regular (and the hypothesis that $s \in C_+$). Thus, by the regularity of μ, there exists $U \subset U_0$ such that $\mu(U) > 0$.

Fix such a U, and notice that for K compact,

$$\sup_{K \subset U} \int \psi_{S_Z}(t; U, K)\, \mu(\mathrm{d}t) = \mu(U) > 0$$

so there exists $K_0 \subset U$ such that $\int \psi_{S_Z}(t; U, K_0)\, \mu(\mathrm{d}t) > 0$. Define $f(\cdot; U) : S_Z \to \mathbb{R}$ as $f(t; U) := \psi_{S_Z}(t; U, K_0) =: f_U(t)$

Then, $V(f_U) = \int f(t; U)\, \mu(\mathrm{d}t) > 0$. Since $f_U \geq 0$, $f_U + r\mathbf{1} \geq r\mathbf{1}$, so that $V(f_U + r\mathbf{1}) = V(f_U) + V(r\mathbf{1}) > V(r\mathbf{1})$ for all $r > 0$. Since f_U is smooth, by Lemma 1 there exists $r > 0$ such that $f_U + r\mathbf{1}$ is the support function of a convex body. Then, $V(f_U + r\mathbf{1}) = V(f_U) + V(r\mathbf{1}) > V(r\mathbf{1})$. Thus, following the arguments above in the example, we see that s is positive. □

2.6 Some comments

Before we end, it is useful to reflect on some of the choices that have been made, principally, in our definition of the sign of a state. Notice that with DLR's definition (which we have adopted), the sign has behavioural implications. For instance, a preference is monotone (in the sense that $A \supset B$ implies $A \succcurlyeq B$) if and only if every subjective state is positive and there is no negative state.

It has been suggested to us that we could avoid the examples presented in this paper by adopting the following classification of the sign of a state. Given a signed measure μ with Hahn–Jordan decomposition $\mu := \mu^+ - \mu^-$, we could say that a state s is *strongly positive* if $\mu^+(s) > 0$ or $s \in \operatorname{supp}(\mu^+) \setminus \operatorname{supp}(\mu^-)$, and that a state s is *strongly negative* if $\mu^-(s) > 0$ or $s \in \operatorname{supp}(\mu^-) \setminus \operatorname{supp}(\mu^+)$. Then all states are either strongly positive or strongly negative, save a set of μ-measure zero.

Unfortunately, such an approach lacks an immediate behavioural interpretation. For instance, suppose μ^+ is a measure with all its weight on $\mathbb{Q} \cap [0, 1]$, the rationals in the unit interval (so that $\operatorname{supp}(\mu^+) = [0, 1]$), and μ^- is the Lebesgue measure on $[0, 1]$. Then, every rational $s \in [0, 1]$ is strongly positive, while no state is strongly negative. Nevertheless, the preference functional induced by $\mu := \mu^+ - \mu^-$ is *not* monotone. (This follows from Theorem 6 above.)

Other such alternative definitions are also possible, but all seem to suffer from similar flaws. This suggests that the definition of the sign of a state proposed by DLR is the appropriate one, since it captures certain desirable behavioural features, and that the issues identified here are substantive.

3 Conclusion

Subjective state models have been extremely useful in capturing hitherto unreachable psychological phenomena such as temptation. A central behavioural property of the subjective state space model is the sign of a state. This sign (of a state) has been used to interpret utility representations as capturing various aspects of the agent's psychological makeup, with a positive sign indicating a preference for flexibility and a negative sign indicating a preference for commitment. Finite state spaces have the attractive property that a state is either positive or negative but never both. In this paper, we show that while finite models can approximate infinite models in utility terms, they cannot capture all the behavioural properties of the infinite state model, such as the sign of a state, in that there are infinite models with states that have both positive and negative sign. We demonstrate this via an example and our methods are useful, more generally, in constructing utility functionals with desired local properties.

References

Dekel, E., Lipman, B., Rustichini, A.: Representing preferences with a unique subjective state space. Econometrica **69**(4), 891–934 (2001)

Dekel, E., Lipman, B., Rustichini, A., Sarver, T.: Representing preferences with a unique subjective state space: a corrigendum. Econometrica **75**(2), 591–600 (2007)

Dekel, E., Lipman, B., Rustichini, A.: Temptation driven preferences. Rev Econ Stud **76**(3), 937–971 (2009)

Dudley, R.M.: Real Analysis and Probability, 2nd edn. Cambridge: Cambridge University Press (2003)

Epstein, L: An axiomatic model of non-Bayesian updating. Rev Econ Stud **73**, 413–436 (2006)

Gul, F., Pesendorfer, W.: Temptation and self-control. Econometrica **69**, 1403–1435 (2001)

Kopylov, I.: Finite additive utility representations for preferences over menus. J Econ Theory **144**(1), 354–374 (2009)

Lee, J.M.: Introduction to Smooth Manifolds. New York: Springer (2003)

Sagi, J.: What is an endogenous state space. Econ Theory **27**, 305–320 (2006)

Sarver, T.: Anticipating regret: why fewer options may be better. Econometrica **76**(2), 263–305 (2008)

Schneider, R.: Convex Bodies: The Brunn–Minkowski Theory. Cambridge: Cambridge University Press (1993)