

On Concave Functions over Lotteries*

Roberto Corrao[†]
MIT

Drew Fudenberg[‡]
MIT

David K. Levine[§]
EUI and WUSTL

November 27, 2023

1 Introduction

This note discusses functions over lotteries that are concave and continuous, but are not necessarily superdifferentiable. Maccheroni [2002]’s Theorem 1 claims that if a function over lotteries is concave, continuous, and satisfies best-outcome independence, it can be written as the minimum of affine functions, and Machina [1984] claims this is true even without the best-outcome-independence condition. However, Section 3 gives an example of a concave and continuous function that satisfies best-outcome independence but cannot be written as the minimum of affine functions, because there is no tangent hyperplane that dominates the functions at the boundary.¹

Section 4 reviews the fact that concavity and upper semi-continuity are equivalent to a representation as the infimum of affine functions, and then shows that these assumptions imply continuity for functions on finite-dimensional lotteries. Therefore, in finite-dimensional simplices, concavity and continuity are equivalent to the “infimum”

*We thank Simone Cerreia-Vioglio, Alex Frankel, Emir Kamenica, Xiaosheng Mu, and Luciano Pomatto for helpful comments, and the Gordon B. Pye Dissertation Fellowship, NSF grant SES 1951056, MIUR (PRIN 20103S5RN3) and Leverhulme Trust for financial support.

[†]Department of Economics, MIT, rcorrao@mit.edu

[‡]Department of Economics, MIT, drew.fudenberg@gmail.com, 50 Memorial Drive, Cambridge, MA 02141

[§]Department of Economics, EUI and WUSTL, david@dklevine.com

¹The example uses a finite-dimensional probability simplex; the issue with boundary points is even more important in the infinite-dimensional setting.

representation.² The “minimum” representation is equivalent to the existence of local utilities (i.e., supporting affine functions) at every lottery, a property that is equivalent to superdifferentiability.³

2 Preliminaries

We study concave functions V on the space \mathcal{F} of probability measures on a compact metric space X , where we identify x with the Dirac measure δ_x . Let $C(X)$ denote the set of continuous functions over X , and endow X with the Borel sigma-algebra. We give \mathcal{F} the topology of weak convergence, so it is metrizable and compact.

Maccheroni [2002] studies preferences over the set $\Delta_\circ(X) \subseteq \mathcal{F}$ of simple lotteries when there is a *best outcome* x^* , i.e. an x^* such that $x^* \succeq F$ for all simple lotteries $F \neq x^*$. Preferences are said to satisfy *best outcome independence* if for all $F, G \in \Delta_\circ(X)$ and $\alpha \in (0, 1)$, $F \succ G$ if and only if $\alpha F + (1 - \alpha)x^* \succ \alpha G + (1 - \alpha)x^*$, so that the usual independence axiom is satisfied with respect to mixtures with x^* . In this case, if \succeq is represented by a continuous utility function V , then V is affine with respect to convex linear combinations of x^* with any arbitrary lottery F . Theorem 1 of that paper makes the following claim:

Claim: *If a function $V : \Delta_\circ(X) \rightarrow \mathbb{R}$ is concave, continuous, and satisfies best-outcome independence, then it can be written as the minimum of a set of affine continuous functions, that is,*

$$V(F) = \min_{w \in \mathcal{W}} \int w(x) dF(x) \tag{1}$$

for some set $\mathcal{W} \subseteq C(X)$.

We provide a counterexample to this claim.⁴ The example has a utility function over lotteries with three outcomes that is continuous, satisfies best-outcome independence, and can be represented by a concave function. However, the associated preferences

²The main result of this section relies on a version of the Hahn-Banach theorem which, for completeness, we state and prove in the appendix.

³This is also implied by the duality results in Dworzak and Kolotilin [2023].

⁴The error in Maccheroni [2002] is an attempt to apply the Hahn-Banach theorem to functions with extended real values in the main claim within the proof of Lemma 4.

over lotteries cannot be represented as the minimum of affine functions, though they can be represented by the infimum of affine functions as we explain below.

3 A Counterexample

We now construct a counterexample to the claim above. Suppose that X has 3 elements so that $\mathcal{F} = \Delta_o(X)$ is the simplex

$$\Delta = \{(p, q) \in [0, 1]^2 : p + q \leq 1\} \subset \mathbb{R}_+^2.$$

We let (\tilde{p}, \tilde{q}) denote an arbitrary point in \mathbb{R}^2 , while we use (p, q) for an arbitrary point in Δ . For every $(\tilde{p}, \tilde{q}) \in \mathbb{R}_+^2$ we let $(r(\tilde{p}, \tilde{q}), \theta(\tilde{p}, \tilde{q})) \in \mathbb{R}_+ \times [0, \pi/2]$ denote the corresponding polar coordinates, and conversely let $(\tilde{p}(r, \theta), \tilde{q}(r, \theta)) \in \mathbb{R}_+^2$ denote the point in the positive orthant corresponding to the polar coordinates (r, θ) .⁵ These mappings define a bijection from the simplex Δ to the subset of $\mathbb{R}_+ \times [0, \pi/2]$ defined by

$$\mathcal{P}_\Delta = \left\{ (r, \theta) \in \mathbb{R}_+ \times [0, \pi/2] : r \leq \frac{\sin(\pi/4)}{\sin(3\pi/4 - \theta)} \right\}.$$

It is convenient to denote points $(p, q) \in \Delta$ by f and use the notation $(r(f), \theta(f)) \in \mathcal{P}_\Delta$ and $f(r, \theta) \in \Delta$.

We will construct a class of utility functions over the simplex by first constructing one of their indifference curves, specifically the one corresponding to $V(f) = -1$. Toward this goal, consider an arbitrary continuous function $\nu : [0, \pi/2] \rightarrow [0, 1]$ such that $\nu(\theta) > 0$ and $(\nu(\theta), \theta) \in \mathcal{P}_\Delta$ for all $\theta \in [0, \pi/2]$, and define the function $\tilde{V}_\nu : \mathbb{R}_+^2 \rightarrow \mathbb{R}$ as

$$\tilde{V}_\nu(\tilde{p}, \tilde{q}) = -\frac{r(\tilde{p}, \tilde{q})}{\nu(\theta(\tilde{p}, \tilde{q}))}.$$

Because $\nu(\theta) > 0$ for all θ , \tilde{V}_ν is well defined at $(0, 0)$ even if $\theta(0, 0)$ is not uniquely defined. Moreover, \tilde{V}_ν is continuous and positive homogeneous: for all $\gamma \geq 0$ and $f \in \Delta$,

$$\tilde{V}_\nu(\gamma f) = -\frac{r(\gamma f)}{\nu(\theta(\gamma f))} = -\frac{\gamma r(f)}{\nu(\theta(f))} = \gamma \tilde{V}_\nu(f)$$

where the second equality follows from the zero-homogeneity property of the function $\theta(p, q)$. We will interpret the restriction of \tilde{V}_ν to lotteries as a utility function, and

⁵Recall that polar coordinates are given by $r(\tilde{p}, \tilde{q}) = \sqrt{\tilde{p}^2 + \tilde{q}^2}$ and $\theta(\tilde{p}, \tilde{q}) = \tan^{-1}(\tilde{q}/\tilde{p})$.

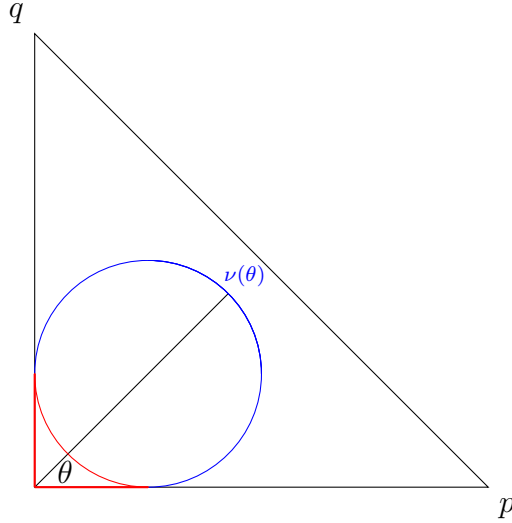


Figure 1: Construction of the indifference curves of the utility function.

denote it by V_ν . Because $r(f) \geq 0$ and $\nu(\theta(f)) > 0$ for all $f \in \Delta$, and $r(f) = 0$ if and only if $f = (0, 0)$, any such utility function is negative. It takes on a strict maximum at the deterministic outcome $x^* = (0, 0)$, which has utility 0, and V_ν assigns utility -1 to each point in the graph of ν , which is the set $\{f(\nu(\theta), \theta) \in \Delta : \theta \in [0, \pi/2]\}$. Note that V_ν is completely determined by the specification of choice of the indifference curve ν .

The intuition behind our counterexample is illustrated by the “snowcone” created by taking a circle tangent to the axes as in Figure 1. For each angle θ , the ray through θ intersects the circle at most twice; we take $\nu(\theta)$ equal to the length of the ray from the origin to the farther point of intersection, that is the one in the upper part of the circle so that the indifference curve ν corresponds to the arc of the circle joining the two tangent points (see the blue arc in Figure 1). Note that this indifference curve is tangent to both the axes.

This diagram corresponds to indifference curves parametrized by $\rho \in (0, 1/4)$ where $\nu_\rho(\theta) = \rho(\cos \theta + \sin \theta + \sqrt{\sin 2\theta})$, which is the distance between the origin and the point in the circle of radius ρ that lies inside the simplex and is tangent to the points $(\rho, 0)$ and $(0, \rho)$. With an abuse of notation, we let V_ρ denote the utility function over the simplex induced by the indifference curve ν_ρ for $\rho \in (0, 1/4)$:

$$V_\rho(f) = -\frac{r(f)}{\rho(\cos \theta(f) + \sin \theta(f) + \sqrt{\sin 2\theta(f)})}.$$

We say that the indifference curve ν is convex if the utility function V_ν constructed as above is such that the utility of any convex combination of two points on the indifference curve lies below the indifference curve: For all f, \tilde{f} such that $V_\nu(f) = V_\nu(\tilde{f}) = -1$, it holds $V_\nu(\lambda f + (1 - \lambda)\tilde{f}) \geq -1$. Notice that the “snowcone” indifference curve in Figure 1 is convex because the unit sphere is a convex set.

Proposition 1. *For all $\rho \in (0, 1/4)$ the indifference curve ν_ρ is convex and V_ρ is concave.*

Concavity of V_ρ follows from the convexity of the indifference curve ν_ρ which can be extended to the other indifference curves by positive homogeneity, as shown in Appendix A.

Next we show that for every $\rho \in (0, 1/4)$, the utility function V_ρ satisfies best-outcome independence with respect to $x^* = (0, 0)$. To see this, consider the lottery $(1 - \alpha)f + \alpha x^*$ for some $f \in \Delta$ and any $\alpha \in (0, 1)$. By homogeneity $V_\rho((1 - \alpha)f + \alpha x^*) = (1 - \alpha)V_\rho(f)$, so f is preferred to \tilde{f} if and only if $(1 - \alpha)f + \alpha x^*$ is preferred to $(1 - \alpha)\tilde{f} + \alpha x^*$.

We finally show that there exists a lottery f such that there cannot be an affine function L on Δ that attains the minimum in Equation 1 for V_ρ . Let S_ρ be the circle with radius ρ and centered at (ρ, ρ) . This circle is tangent to the axes $\tilde{p} = 0$ and $\tilde{q} = 0$ at the points $(\rho, 0)$ and $(0, \rho)$ and is defined by

$$S_\rho = \{(p, q) \in \Delta : (p - \rho)^2 + (q - \rho)^2 \leq \rho^2\}.$$

For every radius $\rho \in (0, 1/4)$, the indifference curve of points $(p, q) \in \Delta$ such that $V_\rho(p, q) = -1$ is the arc from $(\rho, 0)$ to $(0, \rho)$ of S_ρ .⁶

Now fix $\rho \in (0, 1/4)$ and suppose there is a linear function $L(p, q)$ over Δ such that $L(p, q) \geq V_\rho(p, q)$ for all $(p, q) \in \Delta$, with equality at $f = (\rho, 0)$. Because S_ρ is convex, contained in Δ , and such that $(\rho, 0) \in S_\rho$, the supporting hyperplane theorem (see e.g. Theorem 11.6 in Rockafellar [1970]) implies that the set

$$H = \{(\tilde{p}, \tilde{q}) \in \mathbb{R}^2 : L(\tilde{p}, \tilde{q}) = L(\rho, 0)\}$$

is a supporting hyperplane of S_ρ at $(\rho, 0)$. Because S_ρ is a circle, each of its boundary points has a unique supporting hyperplane, so H must coincide with the only support-

⁶In the example plotted in Figure 1, this is the blue arc.

ing hyperplane of S_ρ at $(\rho, 0)$, that is the axis $\tilde{q} = 0$. We then obtain the contradiction

$$L(0, 0) = L(\rho, 0) = V_\rho(\rho, 0) = -1 < 0 = V_\rho(0, 0) \leq L(0, 0)$$

where the first equality follows from the fact that $(0, 0) \in H$ and the second equality and the last weak inequality both follow from the assumptions on L . Therefore, there cannot be a linear function L that dominates V_ρ everywhere and coincides with it at $(\rho, 0)$.⁷ Overall the properties of the utility function V_ρ directly contradict the claim in Section 2.

Machina [1984] analyzes preferences over lotteries $\mathcal{F} = \Delta([0, 1])$ that carry delayed risk. Concretely, consider an agent choosing first $F \in \mathcal{F}$ and then, before the outcome from F has been realized, an action $y \in Y$ from a set of feasible actions. Even if the agent has expected utility preferences over pairs of outcomes and actions, the induced preferences over lotteries is

$$V(F) = \max_{y \in Y} \int u(x, y) dF(x) \tag{2}$$

where u is the utility of the agent and where we assume that the maximum is attained.

Theorem 2 in Machina [1984] states a converse of this fact, that is, if $V : \mathcal{F} \rightarrow \mathbb{R}$ is continuous and convex, then there exists a space of actions Y and a utility function $u(x, y)$ such that V can be represented as in equation 2 with the maximum being attained for every F . In particular, the set of actions is

$$Y = \left\{ y \in C([0, 1]) : \forall F \in \mathcal{F}, V(F) \geq \int y(x) dF(X) \right\}.$$

and $u(x, y) = w(y)$. However, the utility function $-V(F)$, where V is defined as in the snowcone example above, is continuous and convex, but does not have a representation in the form of a maximum as Machina asserts.⁸

⁷A completely symmetric conclusion can be reached for the point $(0, \rho)$.

⁸Theorem 2 of Frankel and Kamenica [2019] asserts that when X is finite, a continuous and concave function H over \mathcal{F} admits a “minimum” representation resembling that of Equation 1, but in that paper the minimum is the same as a supremum, as it is a minimum over a set of linear functions that can take extended real values.

4 Concave Functions and Adversarial Representations

4.1 The Adversarial Representation

We say that V has an *adversarial representation* if

$$V(F) = \inf_{y \in Y} \int u(x, y) dF(x)$$

where Y is a separable metric space and, for every $y \in Y$, $u(\cdot, y)$ is continuous over X .

Theorem 1. *V has an adversarial representation if and only if it is upper semi-continuous and concave.*

The result does not hold with the inf replaced by a minimum, because there may be no tangent hyperplanes at the boundary points of the simplex.⁹ The idea of the theorem is that we can fix this by taking separating hyperplanes that aren't tangent to the concave function, but pass through a point above and near it. Where the function has infinite slope, as we take closer points we get steeper separating hyperplanes, which is why we must use the inf rather than the min. There are various ways to prove this, we provide one in the appendix based on the separating hyperplane theorem.

4.2 Concavity and continuity over finite-dimensional simplices

Notice that $V(F)$ arising from an adversarial representation need not be continuous: Theorem 1 only delivers concavity and upper semi-continuity. Moreover, even in finite-dimensional spaces, there are concave and positive homogeneous functions that fail to be lower semi-continuous, as shown by Example 1 in Appendix C which is derived from an example in Rockafellar [1970] Chapter 10.

However, in studying preference over lotteries, the convex sets on which utility is defined are typically taken to be probability simplices, and the restriction of concave and upper semi-continuous functions to a finite-dimensional subset of \mathcal{F} is continuous,

⁹This problem is especially pervasive in the infinite-dimensional case, where the set of Borel probability measures over a compact metric space has empty (relative) interior when endowed with the topology of weak convergence.

as we show below. Thus the restriction of V to the space of lotteries over a finite $X_0 \subset X$ is continuous.

Theorem 2. *Consider the space $\overline{\mathcal{F}}$ of all convex combinations of N lotteries $\overline{F}^1, \dots, \overline{F}^N$, and suppose that $V(F)$ is concave. Then $V(F)$ restricted to $\overline{\mathcal{F}}$ is lower semi-continuous, and in particular if $V(F)$ is upper semi-continuous then it is continuous.*

When the probability distributions $\overline{F}^1, \dots, \overline{F}^N$ coincide with the point-mass measures over N outcomes, the set $\overline{\mathcal{F}}$ is a probability simplex. The result holds for “generalized simplicies” that are formed by linearly combining N arbitrary lotteries.

Proof. (Adapted from Chapter 10 in Rockafellar [1970].) There is a subset of $\{\overline{F}^1, \dots, \overline{F}^N\}$ that consists of extremal points and whose convex hull is equal to $\overline{\mathcal{F}}$, so w.l.o.g. can assume that $\overline{F}^1, \dots, \overline{F}^N$ are extremal and in particular affinely independent. Hence we may think of points being identified with vectors p on the n -dimensional simplex and we write \overline{p}^i for the basis vectors.

Now consider $\tilde{p} \in \overline{\mathcal{F}}$. Our goal is to prove that for any sequence $p^n \rightarrow \tilde{p}$ we have $\liminf V(p^n) \leq V(\tilde{p})$. Suppose \tilde{p} is not extremal and consider some particular p^n and define $\lambda \equiv \max\{\lambda \geq 0 \mid \lambda \tilde{p} \leq p^n\}$. If $\lambda = 0$ choose i such that $p_i^n = 0$ and $\tilde{p}_i > 0$ otherwise choose i such that $\lambda \tilde{p}_i = p_i^n$ and $\tilde{p}_i > 0$. Consider then the set $\{\overline{p}^1, \dots, \overline{p}^N, \tilde{p}\} - \overline{p}_i$. Since $\tilde{p}_i > 0$ this set is affinely independent. We claim in addition that p^n is a convex combination of these vectors. If $\lambda = 0$ since $p_i^n = 0$ we have p^n already a convex combination of $\{\overline{p}^1, \dots, \overline{p}^N\} - \overline{p}_i$. Otherwise since $p_j^n - \lambda \tilde{p}_j \geq 0$ we may write $p^n = \lambda \tilde{p} + \sum_j (p_j^n - \lambda \tilde{p}_j) \overline{p}^j$ since this is the same as $p_j^n = \lambda \tilde{p}_j + (p_j^n - \lambda \tilde{p}_j)$.

Consider affinely independent sets of the form $(\tilde{p}, \tilde{p}^1, \dots, \tilde{p}^{n-1})$. We showed that if \tilde{p} is not extremal then p^n there exists a set of this form, so p^n lies in the convex hull of the set, and if \tilde{p} is extremal this is true by taking the remaining $n - 1$ vectors to be the remaining basis vectors. Since there are at most n sets of this form it follows that there is a subsequence p^m converging to \tilde{p} that lies entirely in such a set. Clearly $\liminf V(p^n) \leq \liminf V(p^m)$, so it suffices to prove $\liminf V(p^m) \leq V(\tilde{p})$. Because we can write $p^m = \gamma^m \tilde{p} + \sum_{i=1}^{n-1} \gamma_i^m \tilde{p}^i$ with $\gamma^m \rightarrow 1, \gamma_k^m \rightarrow 0$, we have

$$V(p^m) = V(\gamma^m \tilde{p}) + \sum_{i=1}^{n-1} \gamma_i^m V(\tilde{p}^i) \geq \gamma^m V(\tilde{p}) + \sum_{i=1}^{n-1} \gamma_i^m V(\tilde{p}^i) \rightarrow V(\tilde{p})$$

which was our goal. □

4.3 Local utilities and minima

We say that a continuous function $w(x)$ on a compact set X is a *local utility function* for V at $F \in \mathcal{F}$ if $\int w(x)d\tilde{F}(x) \geq V(\tilde{F})$ for all $\tilde{F} \in \mathcal{F}$ and $\int w(x)dF(x) = V(F)$. We say that V has a local expected utility if it has a local utility function at each $F \in \mathcal{F}$.¹⁰ When V has a local expected utility, we denote the set of local expected utilities of V at F by $\mathcal{W}_V(F) \subseteq C(X)$, and set $\mathcal{W}_V = \bigcup_{F \in \mathcal{F}} \mathcal{W}_V(F)$.

Proposition 2. *V has a local expected utility if and only if there exists a set $\mathcal{W} \subseteq C(X)$ such that $V(F) = \min_{w \in \mathcal{W}} \int w(x)dF(x)$. In this case, one such set is $\mathcal{W} = \mathcal{W}_V$.*

The proof of this result is simple and relegated to Appendix A. It is not hard to verify that V has a local utility at F if and only if it is superdifferentiable at F .¹¹ Thus Proposition 2 implies that the “minimum” representation is equivalent to superdifferentiability, as also shown in Dworczak and Kolotilin [2023]. We do not know of a characterization of superdifferentiability of V in terms of more primitive functional conditions or axioms, but there are stronger conditions and axioms that imply superdifferentiability and the “minimum” representation. Notable examples of this are Chatterjee and Krishna [2011], Evren [2014], Sarver [2018], and Ke and Zhang [2020].

References

- Aliprantis, C. D. and K. C. Border (2006). *Infinite Dimensional Analysis*. Springer.
- Bogachev, V. I. and O. G. Smolyanov (2017). *Topological vector spaces and their applications*. Springer.
- Chatterjee, K. and R. V. Krishna (2011). “A nonsmooth approach to nonexpected utility theory under risk.” *Mathematical Social Sciences*, 62, 166–175.
- Corrao, R., D. Fudenberg, and D. K. Levine (2023). “Randomization, surprise, and adversarial forecasters.” [https://www.dropbox.com/scl/fi/zzwv8p9szi4ar8who4kkz/Risk_and_surprise-2023-07-21_online.pdf?rlkey=m.x.s.w.s.v.j.n.k.8.g.e.g.p.0.x.i.y.c.4.b.b.j.a.k.d.l](https://www.dropbox.com/scl/fi/zzwv8p9szi4ar8who4kkz/Risk_and_surprise-2023-07-21_online.pdf?rlkey=m.xs.w.s.v.j.n.k.8.g.e.g.p.0.x.i.y.c.4.b.b.j.a.k.d.l) = 0.

¹⁰If V has a local expected utility, it is concave (see Corrao, Fudenberg, and Levine [2023]).

¹¹That is, there exists a linear function L such that $V(\tilde{F}) \leq V(F) + L(\tilde{F} - F)$ for all $\tilde{F} \in \mathcal{F}$. This differs from the standard definition of superdifferentiability because it only considers probability measures $\tilde{F} \in \mathcal{F}$ instead of all signed measures, but it is the notion used in e.g. Chatterjee and Krishna [2011] and Dworczak and Kolotilin [2023].

- Dworczak, P. and A. Kolotilin (2023). “The persuasion duality.” *Available at SSRN 3474376*.
- Evren, Ö. (2014). “Scalarization methods and expected multi-utility representations.” *Journal of Economic Theory*, 151, 30–63.
- Frankel, A. and E. Kamenica (2019). “Quantifying information and uncertainty.” *American Economic Review*, 109, 3650–80.
- Ke, S. and Q. Zhang (2020). “Randomization and ambiguity aversion.” *Econometrica*, 88, 1159–1195.
- Maccheroni, F. (2002). “Maxmin under risk.” *Economic Theory*, 19, 823–831.
- Machina, M. J. (1984). “Temporal risk and the nature of induced preferences.” *Journal of Economic Theory*, 33, 199–231.
- Rockafellar, R. T. (1970). *Convex analysis*. Vol. 18. Princeton university press.
- Sarver, T. (2018). “Dynamic mixture-averse preferences.” *Econometrica*, 86, 1347–1382.

A Appendix: Proofs

Proposition 1. *For all $\rho \in (0, 1/4)$ the indifference curve ν_ρ is convex and V_ρ is concave.*

Proof of Proposition 1. Fix $\rho \in (0, 1/4)$ and consider the induced indifference curve ν_ρ and utility V_ρ . Suppose $f, \tilde{f} \in \Delta$ are such that $V_\rho(f) = V_\rho(\tilde{f}) = -1$, fix any $\lambda \in [0, 1]$, and define $f_\lambda = \lambda f + (1 - \lambda)\tilde{f}$. Clearly, if $\lambda \in \{0, 1\}$, then $V_\rho(f_\lambda) = -1$. If $\lambda \in (0, 1)$, then f_λ lies strictly below the indifference curve of points such that $V_\rho(f) = -1$, that is the set of points $\Delta_{-1} = \{f \in \Delta : \nu_\rho(\theta(f)) = r(f)\}$. Next, consider the ray passing through the origin $f_0 = (0, 0)$ and f_λ . Let \hat{f} denote the unique point such that this ray intersects Δ_{-1} . By construction, $V_\rho(\hat{f}) = -1$ and $f_\lambda = \gamma \hat{f}$ for some $\gamma < 1$. Finally, by positive homogeneity, $V_\rho(f_\lambda) = V_\rho(\gamma \hat{f}) = \gamma V_\rho(\hat{f}) = -\gamma \geq -1$, as desired.

Next, we show that V_ρ is concave. First, we show that, for all $f \in \Delta$, it holds $\frac{1}{-V_\rho(f)}f \in \Delta$ and $V_\rho\left(\frac{1}{-V_\rho(f)}f\right) = -1$. Indeed, $r\left(\frac{1}{-V_\rho(f)}f\right) = \frac{1}{-V_\rho(f)}r(f) = \nu(\theta(f))$ and $\theta\left(\frac{1}{-V_\rho(f)}f\right) = \theta(f)$, implying that $\left(r\left(\frac{1}{-V_\rho(f)}f\right), \theta\left(\frac{1}{-V_\rho(f)}f\right)\right) \in \mathcal{P}_\Delta$ by the properties

of ν , and hence that $\frac{1}{-V_\rho(f)}f \in \Delta$. Moreover,

$$V_\rho\left(\frac{f}{-V_\rho(f)}\right) = V_\rho\left(\frac{f\nu(\theta(f))}{r(f)}\right) = \frac{\nu(\theta(f))}{r(f)}V_\rho(f) = -\frac{\nu(\theta(f))}{r(f)}\frac{r(f)}{\nu(\theta(f))} = -1,$$

yielding the second part of the claim. Second, observe that for all $\gamma \in [0, 1]$ and $f, \tilde{f} \in \Delta$ we have

$$V_\rho\left(\gamma\frac{f}{-V_\rho(f)} + (1-\gamma)\frac{\tilde{f}}{-V_\rho(\tilde{f})}\right) \geq -1$$

by the first claim and the convexity of the indifference curve $\nu(\theta)$. Third, fix $\lambda \in [0, 1]$, and $f, \tilde{f} \in \Delta$, and define

$$\gamma = \frac{\lambda V_\rho(f)}{\lambda V_\rho(f) + (1-\lambda)V_\rho(\tilde{f})}$$

and observe that

$$1 - \gamma = \frac{(1-\lambda)V_\rho(\tilde{f})}{\lambda V_\rho(f) + (1-\lambda)V_\rho(\tilde{f})}$$

and that both γ and $(1-\gamma)$ are in $[0, 1]$ since $0 \leq \lambda \leq 1$ and $V_\rho \leq 0$. Then

$$\begin{aligned} -1 &\leq V_\rho\left(\gamma\frac{f}{-V_\rho(f)} + (1-\gamma)\frac{\tilde{f}}{-V_\rho(\tilde{f})}\right) = V_\rho\left(-\frac{\lambda f + (1-\lambda)\tilde{f}}{\lambda V_\rho(f) + (1-\lambda)V_\rho(\tilde{f})}\right) \\ &= -\frac{1}{\lambda V_\rho(f) + (1-\lambda)V_\rho(\tilde{f})}V_\rho(\lambda f + (1-\lambda)\tilde{f}). \end{aligned}$$

Because $\lambda V_\rho(f) + (1-\lambda)V_\rho(\tilde{f})$ is negative, $V_\rho(\lambda f + (1-\lambda)\tilde{f}) \geq \lambda V_\rho(f) + (1-\lambda)V_\rho(\tilde{f})$, yielding concavity of V_ρ . \square

Proposition 2. *V has a local expected utility if and only if there exists a set $\mathcal{W} \subseteq C(X)$ such that $V(F) = \min_{w \in \mathcal{W}} \int w(x)dF(x)$. In this case, one such set is $\mathcal{W} = \mathcal{W}_V$.*

Proof of Proposition 2. If V has a local expected utility, then for each $F \in \mathcal{F}$ $\int \hat{w}(x)dF(x) = V(F)$ for all $\hat{w} \in \mathcal{W}_V(F) \subseteq \mathcal{W}_V$, and $\inf_{w \in \mathcal{W}_V} \int w(x)dF(x) \geq V(F)$, so the “only if” part follows. Conversely, let V be such that $V(F) = \min_{w \in \mathcal{W}} \int w(x)dF(x)$ for some set $\mathcal{W} \subseteq C(X)$. Because the minimum is attained, for every F , there exists $w_F \in \mathcal{W}$ such that $V(F) = \min_{w \in \mathcal{W}} \int w(x)dF(x) = \int w_F(x)dF(x)$, so that $w_F \in \mathcal{W}_V(F) \neq \emptyset$. This in turn implies that V has a local expected utility. \square

B Appendix: The Separating Hyperplane Theorem and the Proof of Theorem 1

Inferring properties such as differentiability and concavity from a utility function rests on the separating hyperplane theorem, and one source of error has been misapplying the theorem to infinite-dimensional lotteries. Here we give a careful proof of the separating hyperplane theorem that applies in this setting. Our starting point is the Hahn decomposition theorem as stated in Aliprantis and Border [2006]. For ease of reference, we state that result in the form in which we use it.

Aliprantis and Border [2006] Theorem 5.79. *If the hypograph of $V(F)$, that is the set in $\mathbb{R} \times \mathcal{F}$ consisting of $L = \{(v, F) \in \mathbb{R} \times \mathcal{F} : v \leq V(F)\}$, is closed and convex, then for each singleton set $\{(v, F)\}$ with $v > V(F)$ there is a continuous linear functional separating v from F . This means that there are numbers c_0, z and a continuous function $w_1(x)$ such that for $\tilde{v}, \tilde{F} \in L$ we have $c_0\tilde{v} + \int w_1(x)d\tilde{F}(x) < z$ and $c_0v + \int w_1(x)dF(x) > z$.¹²*

Theorem 3. *Fix $F \in \mathcal{F}$ and a concave and continuous utility function V . For each $v > V(F)$ there exists a continuous function $w(x)$ such that $v \geq \int w(x)d\tilde{F} > V(\tilde{F})$ for all $\tilde{F} \in \mathcal{F}$.*

Proof. We analyze the space of signed measures $H \in \mathcal{M}$ on the Borel σ -sigma algebra of X . The Hahn decomposition theorem says that for any signed measure H the space X can be partitioned into two Borel sets A, B such that for Borel $E \subseteq A$ we have $H(E) \geq 0$ and for $E \subseteq B$ we have $H(E) \leq 0$. The Jordan decomposition further states that there are two positive (ordinary measures) H^+, H^- (uniquely defined) such that for any Hahn decomposition $H^+(B) = 0, H^-(A) = 0$ and $H = H^+ - H^-$. With this in mind, for any continuous function $w : X \rightarrow \mathbb{R}$ and any signed measure we may define $\int w(x)dH(s) = \int w(x)dH^+(x) - \int w(x)dH^-(s)$ where these are ordinary integrals with respect to a signed measure. We may define the total variation $|H| = \int dH^+(x) + \int dH^-(x)$.

Denote the space of bounded continuous functions in the sup norm on X by $C(X)$. On $\mathcal{M} \times C(X)$ we define the operation $\langle H, w \rangle \equiv \int w(x)dH(x)$. If this is continuous and linear in each argument and $\langle H, w \rangle = 0$ for all $w \in C(X)$ if and only if $H = 0$

¹²Crucially, this proposition is false if V can take extended real values, as shown by Bogachev and Smolyanov [2017].

and for all $H \in \mathcal{M}$ if and only if $w = 0$ then $\mathcal{M}, C(X)$ is a dual pair. Continuity follows immediately from $\langle H, w \rangle = \int w(x) dH(x) \leq \|w\| \|H\|$ and this implies $\langle H, w \rangle$ is jointly continuous in the product topology on $\mathcal{M} \times C(X)$.

It follows that \mathcal{H} is a locally convex topological space in the weak topology induced by $C(X)$, and that its continuous linear functionals have the form $\int c(x) dH(x)$ for $c \in C(X)$. Hence this topology relativizes to the subset of probability measures as the topology of weak convergence. That $\langle H, w \rangle = 0$ for all H if $w = 0$ is obvious and only if $w = 0$ follows from considering that the Dirac delta functions δ_x are in \mathcal{H} and $\int w(x) d\delta_x(x) = w(x)$. That $\langle H, w \rangle = 0$ for all w if $H = 0$ is obvious but the “only if” requires some work.

For any $H \neq 0$, we want to find a continuous function $w(x)$ such that $\int w(x) dH(x) \neq 0$. Let A, B be a corresponding Hahn partition of X , and write the Jordan decomposition $H = H^+ - H^-$. Assume without loss of generality that $H^+ \neq 0$. Because X is compact, H^+, H^- are regular measures and in particular $H^+(A) = \sup_{K \subset A} H^+(K)$ and $H^-(B) = \sup_{K \subset B} H^-(K)$ where the supremum is over all compact subsets. Hence we can fix a compact $K^+ \subset A$ such that $|H^+(A) - H^+(K^+)| \leq (1/3)H^+(A)$ and a compact $K^- \subset B$ such that $|H^-(A) - H^-(K^-)| \leq (1/3)H^+(A)$. As K^+, K^- are disjoint and X is a metric space and K^+, K^- are closed, we may use Urysohn’s Lemma to find a continuous function $0 \leq w(x) \leq 1$ which is equal to 1 on K^+ and 0 on K^- . Now write the integral

$$\begin{aligned} \int w(x) dH(x) &= \left[\int_{K^+} w(x) dH^+(x) + \int_{X-K^+} w(x) dH^+(x) \right] \\ &\quad - \left[\int_{K^-} w(x) dH^-(x) + \int_{K^+} w(x) dH^-(x) + \int_{X-K^+-K^-} w(x) dH^-(x) \right] \\ &= H^+(K^+) + \int_{X-K^+} w(x) dH^+(x) - \int_{X-K^+-K^-} w(x) dH^-(x) \\ &\geq (2/3)H^+(A) - (1/3)H^+(A) \geq (1/3)H^+(A) > 0. \end{aligned}$$

Since $\mathcal{M}, C(X)$ is a dual pair, \mathcal{M} is locally convex with respect to the weak topology induced by $C(X)$ in the sup norm: relativized to the probability measures this is the same as the topology of weak convergence.

The hypograph of $V(F)$ is closed because V is upper semi-continuous. Hence by Theorem 5.79 in Aliprantis and Border [2006], for each compact (singleton) set $\{(v, F)\}$ with $v > V(F)$ there are numbers c_0, z and a continuous function $w_1(x)$ such that for

$\tilde{v}, \tilde{F} \in L$ we have $c_0\tilde{v} + \int w_1(x)d\tilde{F}(x) < z$ and $c_0v + \int w_1(x)dF(x) > z$. Applying the first to $\tilde{v}, F \in L$ we have $c_0\tilde{v} + \int w_1(x)dF(x) < z$ so that $c_0\tilde{v} < c_0v$ implying that $c_0 > 0$, since $v > \tilde{v}$. Define $w(x) = -(w_1(x) - z)/c_0$. Observing that $(V(\tilde{F}), \tilde{F}) \in L$, the first inequality says $\int w(x)d\tilde{F} > V(\tilde{F})$ for all \tilde{F} while the second implies $v \geq \int w(x)dF(x)$. \square

Theorem 1. *V has an adversarial representation if and only if it is upper semi-continuous and concave.*

Proof of Theorem 1. First, we show that adversarial implies concave. Consider F, \tilde{F} with $0 \leq \lambda \leq 1$ and y^n such that $\int u(x, y^n)dF(x) \rightarrow V(\lambda F + (1 - \lambda)\tilde{F})$. Consider that $V(F) \leq \int u(x, y^n)dF(x)$ and $V(\tilde{F}) \leq \int u(x, y^n)d\tilde{F}(x)$ so that $\lambda V(F) + (1 - \lambda)V(\tilde{F}) \leq \int u(x, y^n)d(\lambda F + (1 - \lambda)\tilde{F})(x) \rightarrow V(\lambda F + (1 - \lambda)\tilde{F})$.

To show continuity, let $F^n \rightarrow F$ and choose y^m such that $V(F) > \int u(x, y^m)dF(x) - 1/m$. Then

$$V(F^n) \leq \int u(x, y^m)dF^n(x),$$

so

$$\lim V(F^n) \leq \lim \int u(x, y^m)dF^n(x) = \int u(x, y^m)dF(x) < V(F) + 1/m.$$

Hence V is upper semi-continuous.

To prove the other direction, we use Theorem 3 above, which shows that for each $v > V(F)$ there is a continuous function $w(x)$ such that $v \geq \int w(x)d\tilde{F} > V(\tilde{F})$ for all $\tilde{F} \in \mathcal{F}$. Now take Y to be the subset of continuous functions over X for which $\int y(x)d\tilde{F}(x) > V(\tilde{F})$ for all \tilde{F} . This is a separable metric space since it is an open subset of the separable space of all continuous functions endowed with the sup norm. Since for every (v, F) there exists a $y \in Y$ such that $v \geq \int y(x)dF(x) > V(F)$, we see that $V(F) \equiv \inf_{y \in Y} \int u(x, y)dF(x)$. \square

C A concave and positive homogeneous function that is not lower semi-continuous

Example 1. Consider the positive homogeneous function $u(p, q) = -(q + p)^2 p^{-1}$ for $p > 0$. The boundary of the domain is non-linear, which causes a failure of lower

semi-continuity as we next show. This function is concave: we have

$$\begin{aligned}
 Du &= \begin{bmatrix} ((q+p)p^{-1})^2 - 2(q+p)p^{-1} \\ -2(q+p)p^{-1} \end{bmatrix} \\
 D^2u &= \begin{bmatrix} 2((q+p)p^{-1} - 1)(p^{-1} - (q+p)p^{-2}) & \\ & -2p^{-1} \end{bmatrix} \\
 &= \begin{bmatrix} -2p^{-1}((q+p)p^{-1} - 1)^2 & \\ & 2p^{-1}((q+p)p^{-1} - 1) \quad -2p^{-1} \end{bmatrix}
 \end{aligned}$$

which has non-negative diagonal and determinant $\det D^2u = 0$ because u is positive homogeneous so is negative semi-definite. Now restrict this function to the set $0 \leq q \leq \sqrt{p} - p$ on $0 < p \leq 1$. This set is clearly convex. Define $u(0,0) = 0$ and observe that convex combinations of $(1-\lambda)0 + \lambda F$ have $u((1-\lambda)0 + \lambda F) = \lambda u(F) = (1-\lambda)u(0) + \lambda u(F)$ so $u(p,q)$ on $0 \leq q \leq \sqrt{p} - p$ on $0 < p \leq 1$. Consider, however the sequence $F^n = (p^n, \sqrt{p^n} - p^n)$ with $p^n > 0$ and $p^n \rightarrow 0$. Certainly $F^n \rightarrow 0$. On the other hand $u(F^n) = -1$ so the function jumps up in the limit from -1 to 0 hence fails to be lower semi-continuous. \triangle