Journal of Mathematical Economics 52 (2014) 143–147

Contents lists available at ScienceDirect

Journal of Mathematical Economics

journal homepage: www.elsevier.com/locate/jmateco

Expected utility without bounds—A simple proof[☆]

David Dillenberger[a,*], R. Vijay Krishna[b]

[a] Department of Economics, University of Pennsylvania, United States
[b] Duke University, United States

ARTICLE INFO

Article history:
Received 6 May 2013
Received in revised form
17 December 2013
Accepted 31 December 2013
Available online 15 January 2014

Keywords:
Unbounded continuous utility
Weight/growth function

ABSTRACT

We provide a simple proof for the existence of an expected utility representation of a preference relation with an unbounded and continuous utility function.

1. Introduction

Expected utility theory is the cornerstone of most applied work in economics and finance. Many interesting applications of the theory entail working with an unbounded utility on an unbounded domain, together with a rich class of probability distributions. For example, in financial economics, it is customary to posit an exponential (von Neumann–Morgenstern) utility function over monetary prizes of the form $u(x) = -\exp(-\lambda x)$ for some $\lambda > 0$ and normal probability distributions over asset returns. In general, one should be cautious in such situations to make sure that expected utility is well defined. This point is nicely illustrated in the famous St. Petersburg paradox; suppose the utility function over prizes u is unbounded from above. Then, for any natural number n, we can find x_n such that $u(x_n) > 2^n$. Consider such a sequence (x_n) and a probability distribution that yields the prize x_n with probability $1/2^n$. The expected utility of this probability distribution is $+\infty$, which violates the continuity assumption of expected utility. (A similar argument shows that utility functions that are unbounded below are incompatible with any set of lotteries containing all lotteries with countable support.)

A customary way to avoid this problem is to make one of the following choices:

- Restrict attention to simple probability measures, i.e., those with finite support, and allow for unbounded utility that could be discontinuous (or even non-measurable), but still obtain an expectation; or
- Consider all probability measures over $\mathbb{R}$ and strengthen the notion of continuity, but restrict attention to utility functions that are continuous and bounded.

But as Kreps (1988, p. 67) points out (also see Kreps, 2012, p. 100), each of these choices is somewhat 'disappointing', in that they embody drastic tradeoffs. While the first option is very permissive with respect to the utility functions allowed, it rules out probability measures with infinite support (such as the normal distribution over $\mathbb{R}$). Similarly, the second option permits all probability measures, but it rules out unbounded utility functions, such as CARA utility. As exponential utility and a normal distribution of returns is a parametrization that is analytically convenient, it would be desirable to find conditions under which this specification, as well as many other, can be accommodated within the expected utility framework. In particular, this means that in order to work with some class of unbounded utility functions, one must identify both the appropriate class of probability measures to look at and the appropriate topology on the space of these measures.

A more satisfactory solution is provided by Föllmer and Schied (2011, Theorem 2.9) (henceforth FS). They show that the natural space to look at is the space of functions that satisfy a *growth condition* and that the corresponding domain of probability measures is all probability measures that have finite integral with respect to the function governing this growth condition. They then provide a representation with continuous and unbounded utility (but bounded relative to the growth condition). Their proof is essentially 'from scratch' in the sense that they prove their theorem from first principles and is along the lines hinted at by Kreps (2012, p. 100).

In this paper we provide a much simpler proof of FS's result. We exhibit an isometry – details are in Section 2 – between pairs of spaces that allows us to reduce the problem to that of exhibiting

☆ We thank an anonymous referee and an associate editor for helpful comments. Krishna gratefully acknowledges support from the National Science Foundation via grant SES-1132193.

* Corresponding author.
E-mail addresses: ddill@sas.upenn.edu (D. Dillenberger), vijay.krishna@duke.edu (R. Vijay Krishna).

0304-4068/$ – see front matter © 2014 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jmateco.2013.12.011

continuous and bounded utility. Thus, with a simple isometry and a little linear algebra, we can obtain a representation with continuous and unbounded utility.

We note that our proof is, in contrast with the proof of FS, virtually free of calculations. Instead, it leverages the proof of the standard case with bounded utility—see Section 3. In addition to providing a utility representation, we also discuss a preorder on the space of all growth conditions and show that with more stringent growth conditions, fewer probability measures are admissible for analysis.

We provide simple examples with various growth conditions. The examples also emphasise the price of allowing unbounded utility: we can no longer consider *all* probability measures on the domain, and the condition for the convergence of probability measures is correspondingly more stringent.

Before we proceed, it is worthwhile to consider the value of our Theorem 3.1. For instance, it may be argued that in most applied models, if the modeller is considering an unbounded utility function, then he simply considers a particular set of lotteries with finite expected utility (calculated using the utility function in question). While this is undoubtedly true, our approach provides a straightforward way to delineate *all* lotteries with finite expected utility. But more importantly, our approach makes clear the natural topology on this space of admissible lotteries. This latter observation is important because it is central to the analysis of counterfactuals where, for instance, alternative lotteries may be considered that may even be perturbations of the given lottery. Such analysis is feasible only if one understands the topology on the space of admissible lotteries. Put somewhat differently, this last observation implies that in applied models, topologies on the space of (admissible) probability measures should be determined in conjunction with the space of (unbounded) utilities in question.

We note that the topology on the space of admissible lotteries is finer than the topology of weak convergence. It is easy to see that such a property is necessary. If not, we could approximate any lottery with simple lotteries in the topology of weak convergence, which would violate the continuity of preferences.

2. Setting and preliminaries

First we introduce some notation. Let X be a separable metric space, $C(X)$ be the space of continuous functions on X, and $C_b(X)$ be the space of bounded, continuous functions on X. Similarly, $\mathscr{M}(X)$ is the space of finite (Borel) measures on X, $\mathscr{M}^+(X)$ is the positive orthant of $\mathscr{M}(X)$ (i.e., the space of positive measures), and $\mathscr{P}(X)$ is the space of *probability* measures on X. We will endow $\mathscr{M}(X)$ with the weak* topology, via the dual pairing $\langle \mathscr{M}(X), C_b(X)\rangle$. Convergence in this topology is denoted by w^*.

Let $\mathscr{F} := \{g \in [1,\infty)^X : g \in C(X)\}$ be the set of all continuous and positive functions that are bounded away from zero (in particular, for $g \in \mathscr{F}$, $\inf_{x\in X} g(x) \geqslant 1$), with typical members f, g. This will be the set of *weight functions*.

For the rest of the paper, we use the following notational conventions. For a function $f \in \mathbb{R}^X$ and measure μ on X, the integral $\int_X f(x)\mu(\mathrm{d}x)$ is denoted by μf. Similarly, for $g \in \mathscr{F}$, the measure $\nu(\mathrm{d}x) := g(x)\mu(\mathrm{d}x)$ is written as $g\mu$.

Let $g \in \mathscr{F}$ and define $W_g(X)$ to be the space of continuous functions $u \in C(X)$ such that $\sup_{x\in X} |u|/g(x) < \infty$. It is easy to see that $W_g(X)$ is a linear space. In fact, we can make $W_g(X)$ a *normed* space, by defining $\|u\|_{g,\infty} := \sup_{x\in X} |u|/g(x)$, which is easily seen to be a norm. Clearly, $u \in W_g(X)$ if, and only if, $\|u\|_{g,\infty} < \infty$. (As we shall see below, $W_g(X)$ equipped with the norm $\|\cdot\|_{g,\infty}$ is, in fact, a Banach space.) Intuitively, $W_g(X)$ consists of all functions that do not grow faster than g; this is why g is also said to imply a growth condition.

Let $\mathscr{M}_g(X) := \{\mu \in \mathscr{M}(X) : \mu g < \infty\}$ be the space of measures on X whose g-integral is finite. Then, $\mathscr{M}_g^+(X)$ is the positive orthant of $\mathscr{M}_g(X)$, and $\mathscr{P}_g(X)$ is the space of probability measures on X that have finite g-expectation. If we choose g to be a constant function, say $g = \mathbf{1}$, then $\mathscr{P}_{\mathbf{1}}(X) = \mathscr{P}(X)$. Notice that $\mathscr{P}_g(X)$ may be a strict subset of $\mathscr{P}(X)$ when g is not constant. For instance, when $X = \mathbb{R}$, $g = \max[1, |x|]$ implies that $\mathscr{P}_g(\mathbb{R})$ does not contain any probability measure with infinite expectation (like the Cauchy distribution). Notice also that for any $g \in \mathscr{F}$, $\mathscr{P}_g(X)$ is a convex set (and hence a mixture space). To see this, let $\mu_1, \mu_2 \in \mathscr{P}_g(X)$, and let $t \in (0,1)$. Then, $0 \leqslant \int g\,\mathrm{d}\big(t\mu_1 + (1-t)\mu_2\big)(x) = t\int g\,\mathrm{d}\mu_1(x) + (1-t)\int g\,\mathrm{d}\mu_2(x) < \infty$ because both μ_1 and μ_2 have finite g expectation by virtue of being in $\mathscr{P}_g(X)$.

Recall that $ba_n(X)$ the space of finite normal charges on X (endowed with the Borel σ-algebra) is the dual of $C_b(X)$. Our first result describes the dual space of $W_g(X)$.

Proposition 2.1. *The operator $T : W_g(X) \to C_b(X)$ defined as $Tu := u/g$ is a bijective isometry. Therefore, the norm dual of the space $W_g(X)$ is isometrically isomorphic to the space of all normal bounded charges μ on X such that μg is finite.*

Proof. Consider the *map* $T : W_g(X) \to C_b(X)$ given by $Tu := u/g$. It is easy to see the following: (i) T is an isometry, (ii) T is invertible, and (iii) T is linear. Thus, T is an operator and is a bijective isometry. This implies the adjoint of T, namely $T^* : ba_n(X) \to W_g(X)^*$, is a bijective isometry. It is also easy to see that if $p \in ba_n(X)$, $T^*p(\mathrm{d}x) = p(\mathrm{d}x)/g(x)$.

Finally, to see that $W_g(X)^*$ is the space of all normal, bounded charges μ on X with μg finite, let us first assume that μg is not finite (where we assume, without loss of generality, that μ is positive). This implies $g\mu \notin ba_n(X)$, which implies that μ is not the image of any element of $ba_n(X)$, and hence not a member of $W_g(X)^*$. Conversely, let us assume that μg is finite, so $g\mu \in ba_n(X)$. Then, $T^*(g\mu) = \mu \in W_g(X)^*$. □

It follows from the definition of T^* that $T^*\mathscr{M}(X) = \mathscr{M}_g(X)$ and $T^*\mathscr{M}^+(X) = \mathscr{M}_g^+(X)$. Therefore, we will endow $\mathscr{M}_g(X)$ with the weak* topology, via the duality $\langle \mathscr{M}_g(X), W_g(X)\rangle$ and convergence in this topology will be denoted by w_g^*. By the definition of T, for any net (μ_α) in $\mathscr{M}_g^+(X)$, $\mu_\alpha \to_{w_g^*} \mu$ if, and only if, $\mu_\alpha/g \to_{w^*} \mu/g$. Notice also that if $u \in W_g(X)$, then $|u(x)| \leqslant \|u\|_{g,\infty}\, g(x)$, so that for all $\mu \in \mathscr{P}_g(X)$, we have $\mu|u| \leqslant \|u\|_{g,\infty}\, \mu g < \infty$.

We define the binary relation on $\trianglerighteq$ on $\mathscr{F}$ as follows: for, $g, h \in \mathscr{F}$, $g \trianglerighteq h$ if $g \in W_h(X)$. Note that $\trianglerighteq$ is reflexive and transitive, but not antisymmetric. Hence $\trianglerighteq$ is a *preorder*. Intuitively, $g \trianglerighteq h$ if h grows faster than g.

Corollary 2.2. *Let $g \trianglerighteq h$. Then, $u \in W_g(X)$ implies $u \in W_h(X)$ and $\mu \in \mathscr{P}_h(X)$ implies $\mu \in \mathscr{P}_g(X)$.*

Proof. By assumption $g \trianglerighteq h$, so $g \in W_h(X)$. Suppose now that $u \in W_g(X)$, so that

$$\begin{aligned}\sup_{x\in X} |u(x)|/h(x) &= \sup_{x\in X}\left[\frac{|u(x)|}{g(x)}\frac{g(x)}{h(x)}\right]\\ &\leqslant \left[\sup_{x\in X}|u(x)|/g(x)\right]\left[\sup_{x\in X} g(x)/h(x)\right]\\ &< \infty.\end{aligned}$$

Suppose now that $\mu \in \mathscr{P}_h(X)$. Then, because $g \trianglerighteq h$, it follows that there exists $c > 0$ such that $g(x) \leqslant ch(x)$ for all $x \in X$. Thus, $\int g(x)\,\mu(\mathrm{d}x) \leqslant c\int h(x)\,\mu(\mathrm{d}x) < \infty$, from which it follows that $\mu \in \mathscr{P}_g(X)$. □

Corollary 2.3. *$W_g(X) = W_h(X)$ if, and only if, $g \trianglerighteq h$ and $h \trianglerighteq g$.*

Proof. The 'if' part was proved above in Corollary 2.2. To see the 'only if', notice that $g \in W_g(X)$ for all $g \in \mathscr{F}$, so that if $W_g(X) = W_h(X)$, then $g \in W_h(X)$ and $h \in W_g(X)$. But this implies $g \trianglerighteq h$ and $h \trianglerighteq g$, as required. □

Corollary 2.4. *If $\mu \notin \mathscr{P}_g(X)$, then $\mu g = \infty$.*

Proof. If μ is a probability measure on $\mathscr{P}(X)$ such that for all $u \in W_g(X)$ we have $\mu|u| < \infty$, then setting $u = g$ implies $\mu \in \mathscr{P}_g(X)$, which establishes the contrapositive. □

3. Representation

We shall rely on the following result which is essentially a restatement of the standard expected utility theorem for bounded continuous utility—see, for instance, Theorem 3 of Grandmont (1972) or Theorem 5.21 and Corollary 5.22 of Kreps (1988).

Theorem 3.0. *Let $\hat{U} : \mathscr{M}^+(X) \to \mathbb{R}$ be linear and w^*-continuous. Then, there exists $\hat{u} \in C_b(X)$ such that $\hat{U}(p) = \int_X \hat{u}(x)p(\mathrm{d}x)$ for all $p \in \mathscr{M}^+(X)$.*

We are now ready to state our main theorem (which is Theorem 2.9 of FS).

Theorem 3.1. *Let $\succsim \subset \mathscr{P}_g(X) \times \mathscr{P}_g(X)$ be a binary relation. Then, $\succsim$ is a preference relation (i.e., is complete and transitive) that is w_g^*-continuous and satisfies independence if, and only if, there exists a function $u \in W_g(X)$ such that for all $\mu, \nu \in \mathscr{P}_g(X)$, $\mu \succsim \nu$ if, and only if, $\mu u \geqslant \nu u$. Moreover, u is unique up to positive affine transformation.*

Proof. Notice that the uniqueness of the function u follows from the Mixture Space Theorem. Therefore, it suffices to establish existence.

By the Mixture Space Theorem, there exists $U_0 : \mathscr{P}_g(X) \to \mathbb{R}$ that represents $\succsim$, and is linear and w_g^*-continuous. Let $U : \mathscr{M}_g^+(X) \to \mathbb{R}$ denote the (unique) extension of U_0 to $\mathscr{M}_g^+(X)$ by linearity. Notice that U is w_g^*-continuous.

Now define $\hat{U} : \mathscr{M}^+(X) \to \mathbb{R}$ as follows: $\hat{U} := U \circ T^*$, so $\hat{U}(p) = U(p/g)$ where $p \in \mathscr{M}^+(X)$ and T^* is as in Proposition 2.1. The function $\hat{U}$ is linear and continuous (because it is a composition of linear and continuous functions). By Theorem 3.0, there exists $\hat{u} \in C_b(X)$ such that $\hat{U}(p) = \int_X \hat{u}(x)p(\mathrm{d}x)$. In other words, for each $\mu \in \mathscr{M}_g^+(X)$, we have $U(\mu) = \hat{U}(g\mu) = \int_X \hat{u}(x)g(x)\mu(\mathrm{d}x) = \int_X u(x)\mu(\mathrm{d}x)$, where $u(x) := \hat{u}(x)g(x) \in W_g(X)$. Therefore, for all $\mu, \nu \in \mathscr{P}_g(X)$, $\mu \succsim \nu$ if, and only if, $\mu u \geqslant \nu u$, which proves the theorem. □

4. Examples

Example 4.1 (*Risk Neutral Agents*)**.** Let $X := \mathbb{R}$ and $g_1(x) := \max[1, |x|]$. Notice that $u(x) = x$ is now in $W_{g_1}(\mathbb{R})$. In other words, all risk neutral expected utility functions are in $W_{g_1}(\mathbb{R})$. Therefore, if $\mu \in \mathscr{P}_{g_1}(X)$, then μ has finite mean. In fact, if $\mu \in \mathscr{P}(X)$ has finite mean, then $\mu \in \mathscr{P}_{g_1}(X)$. This follows immediately from the observation that $\mu g_1 \leqslant \int_{[-1,1]} 1 \cdot \mu + \int_{\mathbb{R}\setminus[-1,1]} |x|\, \mu \leqslant \mu([-1, 1]) + \mu\, |x| < \infty$.

It is useful to note that the lottery with the Cauchy distribution, which does not have a mean, is not in $\mathscr{P}_{g_1}$ given our particular choice of g_1. Note also that exponential utilities (CARA) are not in $W_{g_1}(\mathbb{R})$. ▲

Next consider the case of CARA utilities.

Example 4.2. Let $X := \mathbb{R}$ and $g_2(x) := \exp(\alpha\, |x|)$ where $\alpha > 0$. We claim that the utility function $u(x) = -\exp(-\beta x)$ is in $W_{g_2}(\mathbb{R})$ for $\beta \in (0, \alpha)$. To see this, notice that

$$\frac{|u(x)|}{g_2(x)} = e^{-\beta x - \alpha|x|} = \begin{cases} e^{-x(\alpha+\beta)} & \text{if } x \geqslant 0 \\ e^{x(\alpha-\beta)} & \text{if } x \leqslant 0. \end{cases}$$

Therefore, $u \in W_{g_2}(\mathbb{R})$ if, and only if, $\alpha \geqslant \beta$. In other words, all CARA utility functions with Arrow–Pratt risk aversion parameter β are in $W_{g_2}(\mathbb{R})$ if, and only if, $\beta \leqslant \alpha$.

In addition, all the normal distributions are in $\mathscr{P}_{g_2}(\mathbb{R})$. To see this, let μ be the lottery with mean 0 and variance σ^2. Then, $\mu g_2 \propto \int_0^\infty \exp(\alpha x) \exp(-x^2/2\sigma^2)\, \mathrm{d}x$. But

$$\int_0^\infty e^{\alpha x - x^2/2\sigma^2}\, \mathrm{d}x = e^{\alpha^2\sigma^2/2} \int_0^\infty e^{-\frac{1}{2\sigma^2}(x-\alpha\sigma^2)^2}\, \mathrm{d}x$$

and the latter is finite, so $\mu \in \mathscr{P}_{g_2}(\mathbb{R})$. It is easy to see that considering normal distributions with zero mean is without loss of generality.

Finally, notice that the lottery on $\mathbb{R}_-$ with exponential distribution $F(x) = \min[1, \exp(\beta x)]$, which has mean $-1/\beta$, is in $\mathscr{P}_{g_2}(\mathbb{R})$ if, and only if, i.e., if, and only if, the mean is sufficiently large. This is because $\int_{\mathbb{R}} g_2(x)\, \mathrm{d}F(x) = \int_{-\infty}^0 \exp((\beta - \alpha)x)\, \mathrm{d}x$, which is finite if, and only if, $\beta > \alpha$. ▲

A drawback of the weight function $g_2(x) = \exp(\alpha\, |x|)$ is that it rules out exponential lotteries with high mean that are supported on $\mathbb{R}_+$. We can rectify this by picking a different weight function.

Example 4.3. Let $X := \mathbb{R}$ and $g_3(x) := \max[\exp(-\alpha x), x]$ where $\alpha > 0$. Once again, the utility function $u(x) = -\exp(-\beta x)$ is in $W_{g_3}(\mathbb{R})$ if, and only if, $\alpha \geqslant \beta$. In other words, all CARA utility functions with Arrow–Pratt risk aversion parameter β are in $W_{g_3}(\mathbb{R})$ if, and only if, $\beta \leqslant \alpha$. As before, all the normal distributions are also in $\mathscr{P}_{g_3}(\mathbb{R})$. Moreover, all risk neutral utilities are in $W_{g_3}(\mathbb{R})$.

Moreover lotteries with distribution $F(x) := 1 - \exp(-\gamma x)$, where $x \geqslant 0$, which have mean $1/\gamma$, are all in $\mathscr{P}_{g_3}(\mathbb{R})$. ▲

The examples above illustrate Corollary 2.2 because $g_3 \trianglerighteq g_2$ and there exist probability measures that are in $W_{g_3}(\mathbb{R})$ but not in $W_{g_2}(\mathbb{R})$. A drawback to Examples 4.2 and 4.3 is that due to our choice of the weight functions g_2 and g_3, $W_{g_2}(\mathbb{R})$ and $W_{g_3}(\mathbb{R})$ do not contain all CARA utility functions even though $\mathscr{P}_{g_2}(\mathbb{R})$ and $\mathscr{P}_{g_3}(\mathbb{R})$ contain all the normal distributions. This is because the weight function $g_2(x) := \exp(\alpha\, |x|)$ where $\alpha > 0$ does not grow sufficiently quickly to ensure that every CARA utility is in $W_{g_2}(\mathbb{R})$. We now consider a weight function that allows for all CARA utilities.

Example 4.4. Let $X := \mathbb{R}$ and let $g_4(x) := \max\big[\exp\big(\exp(-x)\big)\big]$. We can now ensure that every CARA utility function is in $W_{g_4}(X)$. To see this, consider the CARA utility function $-\exp(-\alpha x)$, and notice that $\lim_{x\to-\infty}\big[\exp(-\alpha x)/\exp[e^{-x}]\big] = 0$, from which it follows immediately that $-\exp(-\alpha x) \in W_{g_4}(\mathbb{R})$. However, the lottery with normal distribution (mean r, variance σ^2) is not in $\mathscr{P}_{g_4}(\mathbb{R})$, because e^x grows faster than x^2. ▲

This suggests that we should look for a weight function that grows faster than any linear function but that does not grow faster than x^2.

Example 4.5. Let $X := \mathbb{R}$ and let

$$g_5(x) := \begin{cases} \exp\big((-x)^{1+t}\big) & x < 0 \\ e & x \geqslant 0 \end{cases}$$

where $t \in (0, 1)$. Let $u(x) = -\exp(-\beta x)$ be a CARA utility function where $\beta \geqslant 0$. To prove that $\sup_{x\in X} |u(x)|/g_5(x) < \infty$, it suffices to show that $\lim_{x\to-\infty} \exp\big(-\beta x - (-x)^{1+t}\big) < \infty$. Notice that $\lim_{x\to-\infty} -\beta x - (-x)^{1+t} = \lim_{x\to-\infty}(-x)\big[\beta - (-x)^t\big] = -\infty$ from which it follows immediately that $\lim_{x\to-\infty} \exp\big(-\beta x - (-x)^{1+t}\big) = 0 < \infty$. Thus, every CARA utility function is in $W_{g_5}(\mathbb{R})$.

We shall now show that every normal distribution is in $\mathscr{P}_{g_5}(\mathbb{R})$. Consider a normal distribution with mean r and variance σ^2. Such a distribution is in $\mathscr{P}_{g_5}(\mathbb{R})$ if, and only if, $\int_{\mathbb{R}} g_5(x) \exp\big((x - r)^2/2\sigma^2\big)\, \mathrm{d}x$ is finite. Observe that this expectation is finite if, and only if, $\int_{-\infty}^0 \exp\big((-x)^{1+t}\big) \exp\big((x - r)^2/2\sigma^2\big)\, \mathrm{d}x$ is finite or equivalently, if $\int_0^\infty \exp(x^{1+t}) \exp\big((x - r)^2/2\sigma^2\big)\, \mathrm{d}x$ is finite. The function $x^{1+t} - (x - r)^2/2\sigma^2$ is decreasing on the set $[y^*, \infty)$ for some sufficiently large $y^* > 0$, so to show that $\int_0^\infty \exp(x^{1+t}) \exp\big((x - r)^2/2\sigma^2\big)\, \mathrm{d}x$ is finite, it suffices to show that $\int_{y^*}^\infty \exp(x^{1+t}) \exp\big((x - r)^2/2\sigma^2\big)\, \mathrm{d}x$ is finite.

We have just established that $\exp\big((x - r)^2/2\sigma^2\big)$ is decreasing on the set $[y^*, \infty)$. It is also positive and continuous. Therefore,

$\int_{y^*}^{\infty} \exp(x^{1+t}) \exp\big((x-r)^2/2\sigma^2\big)\,\mathrm{d}x$ is finite if, and only if, $\sum_{n>y^*}^{\infty} \exp(n^{1+t}) \exp\big((n-r)^2/2\sigma^2\big)$ is finite—this is the integral test for the convergence of a series. This series converges if, and only if, $\lim n \to \infty |a_{n+1}/a_n| < 1$ (the ratio test), where $a_n := \exp(n^{1+t}) \exp\big((n-r)^2/2\sigma^2\big)$.

The ratio $a_{n+1}/a_n = \exp\big((n+1)^{1+t} - n^{1+t} - 1/2\sigma^2 - (n-r)/\sigma^2\big)$. Therefore, it suffices to consider the limit of the exponent. Notice that x^{1+t} is a convex function because $t > 0$, so $(n+1)^{1+t} - n^{1+t} \leqslant (1+t)(n+1)^t$. Therefore,

$$\begin{aligned}
&(n+1)^{1+t} - n^{1+t} - 1/2\sigma^2 - (n-r)/\sigma^2 \\
&\quad \leqslant (1+t)(n+1)^t - 1/2\sigma^2 - (n-r)/\sigma^2 \\
&\quad = (1+t)(n+1)^t - (n+1)/\sigma^2 - 1/2\sigma^2 + (1+r)/\sigma^2 \\
&\quad = (n+1)^t\big[(1+t) - (n+1)^{1-t}/\sigma^2\big] - 1/2\sigma^2 + (1+r)/\sigma^2
\end{aligned}$$

and $\lim_{n\to\infty}(n+1)^t\big[(1+t) - (n+1)^{1-t}/\sigma^2\big] = -\infty$, which establishes that $\lim_{n\to\infty} a_{n+1}/a_n = 0$. This proves the claim that every lottery with normal distribution and parameters (r, σ^2) is in $\mathscr{P}_{g_5}(\mathbb{R})$. ▲

A utility function displays *infinite (absolute) risk aversion* if it is more risk averse than every CARA utility function at every wealth level. Notice that the utility function $u(x) = -g_5(x)$ (which is increasing and concave), where g_5 is defined in Example 4.5, does not exhibit infinite risk aversion. Nevertheless, every CARA utility is in $\mathscr{P}_{g_5}(\mathbb{R})$. This is because the weight function g_5 has the following property: for every $\beta \geqslant 0$, there is a wealth level $x_\beta < 0$ such that at this wealth level, the (Arrow–Pratt) risk aversion of $u = -g_5$ is greater than β.

We end this section with a counterexample to a claim of Huang and Litzenberger (1988, p. 14). They claim that if the utility function is concave, then '…expected utilities of [lotteries] having finite expectations will be finite even when [the utility function] is unbounded.'

Example 4.6. Consider the CARA utility function $u : \mathbb{R} \to \mathbb{R}$ given by $u(x) = -e^{-x}$. Let μ be the lottery that is exponentially distributed on $\mathbb{R}_-$. Then, the expected value of the lottery (its mean) is $\int_{\infty}^{0} xe^x\,\mathrm{d}x = [xe^x]_{-\infty}^{0} - \int_{-\infty}^{0} e^x\,\mathrm{d}x = 0 - 0 - 1 = -1$. The expected utility of μ is $\int u(x)\mu(\mathrm{d}x) = \int_{-\infty}^{0} -e^{-x}e^x\,\mathrm{d}x = -\int_{-\infty}^{0}\mathrm{d}x = -\infty$. ▲

Huang and Litzenberger's claim is true (for the utility function above) if we restrict attention to lotteries whose support is bounded below; this follows from Jensen's inequality. More generally, the claim is true if the utility function restricted to the support of the probability measure in question is bounded below. Example 4.6 considers a lottery whose support is unbounded below.

5. Related literature

The question addressed in this paper is well known and has been analysed from various perspectives. In this section, we will place our approach in context.

The expected utility rule was axiomatised by von Neumann and Morgenstern (1944) who considered the special case where the set of probability measures under consideration is $\mathscr{P}_0(X)$, the space of *simple* probability measures on X.[1] It is well known – for instance, from the St. Petersburg paradox described in the Section 1 – that the expected utility rule is incompatible with lotteries that have unbounded support. A specific axiomatisation of expected utility that allows for all discrete probability measures is Blackwell and Girshick (1954), who note that the resulting vN–M utility function over prizes must necessarily be bounded.

Further axiomatic developments include Ledyard (1971), DeGroot (1970), and Fishburn (1975, 1976). These are significant because they derive an expected utility representation while allowing for unbounded (vN–M) utility. Clearly, this means that the domain of probability measures must somehow be restricted. We now provide a brief description of their respective constructions. Let P denote some convex subset of $\mathscr{P}(X)$ such that P contains $\mathscr{P}_0(X)$. Ledyard and DeGroot first consider the space $\mathscr{P}_b(X)$, which consists of probability measures with bounded support, whence the probability of the set $\{x : x_1 \succsim x \succsim x_2\}$ is 1 for some $x_1, x_2 \in X$. (Here, $x_1 \succsim x$ if, and only if, $\delta_{x_1} \succsim \delta_x$, where δ_x is the degenerate lottery (Dirac measure at x) that gives x with probability one.) Standard techniques allow them to find an expected utility representation on this domain, and the resulting utility $u \in \mathbb{R}^X$ need not be bounded. They then extend the representation to the space of all unbounded probability measures that have finite u expectation.

In contrast, Fishburn (1975) considers some fixed convex subset P of $\mathscr{P}(X)$ that contains $\mathscr{P}_0(X)$ and directly imposes conditions on the preference that ensures the existence of an expected utility representation.[2] For simplicity, let us consider the case where $X = \mathbb{R}$, and $\succsim$ restricted to X is (weakly) monotone, i.e., preference over outcomes is monotone. Fishburn (1975) notes that if $\succsim$ is a preference on P and satisfies the usual vN–M axioms, then it has a linear representation, V. We may define the utility function $u \in \mathbb{R}^X$ as $u(x) := V(\delta_x)$ for all $x \in X$. In that case, on the subdomain $\mathscr{P}_0(X)$ of simple probability measures, $V(p) \geqslant V(q)$ if, and only if, $\sum_x u(x)p(x) \geqslant \sum_x u(x)q(x)$, where $p, q \in \mathscr{P}_0(X)$. He then shows (Theorem 1) that if $\succsim$ satisfies an additional dominance axiom, then for all $p \in P$ with supports that are unbounded above but bounded below, it must necessarily be the case that $V(p) \geqslant \int u(x)\,\mathrm{d}p(x)$. Fishburn further introduces another axiom (Axiom 5′) which requires that the upper truncation of p is well behaved. Formally, Axiom 5′ from Fishburn (1975) requires that if p has supported unbounded above but bounded below, and if $p_0 \in \mathscr{P}_0(X)$ is such that $p \succ p_0$, then there is some upper truncation of p that is at least as good as p_0. Clearly, this implies that $V(p) \leqslant \int u(x)\,\mathrm{d}p(x)$, which proves that $V(p) = \int u(x)\,\mathrm{d}p(x)$. Intuitively, the axiom ensures that u is such that u decreases sufficiently quickly so as to ensure that the expected utility is well defined.

There are two main differences between the present work and that of Ledyard (1971), DeGroot (1970), and Fishburn (1975). The first is that these works make no use of the topological structure (except for Fishburn (1975) who requires some mild connectivity assumption) of the prize space. This means that these papers must necessarily be silent about the continuity of the vN–M utility u, even in the case where the prize space $X = \mathbb{R}$. In contrast, our approach immediately delivers a continuous vN–M utility function while only requiring that the prize space X be a separable metric space.

The second difference is that the above works implicitly recover a growth condition on the utility function (via the preference ordering on the set P), while we explicitly impose a growth condition on the space of probability measures, which immediately leads to a natural dual space of utilities satisfying a growth condition. Put differently, the essence of the problem boils down to requiring that (i) the vN–M utility function not grow too quickly, and (ii) the tail probabilities decrease sufficiently fast. The properties

[1] A generalisation of von Neumann and Morgenstern's theorem is provided by Herstein and Milnor (1953), who provide necessary and sufficient conditions for the existence of a linear representation – for instance, a function $V : \mathscr{P}(X) \to \mathbb{R}$ such that $p \succsim q$ if, and only if, $V(p) \geqslant V(q)$ – on a *mixture space*, which generalises the notion of a convex set.

[2] Fishburn (1976) provides a specialisation of Fishburn (1975) for the case where the set of prizes is $X := \mathbb{R}_+$.

of the tail probabilities are determined by the set P under consideration. (This is because for any convex set $P \subset \mathscr{P}(X)$, there exists a weight function g such that $P \subset \mathscr{P}_g(X) \subset \mathscr{P}(X)$.) Thus, the set P automatically implies a growth condition on tail probabilities. If expected utilities are to be well defined, then there must be a corresponding growth condition on the vN–M utility function, and the various approaches discussed above achieve this in different ways.

In contrast, our approach fixes a growth condition (implied by a weight function) such that relative to this weight function (i′) the only relevant vN–M utility functions are those that do not grow faster than the weight function, and (ii′) the only relevant probability measures are those whose tail probabilities diminish sufficiently quickly so as to render the weight function integrable. Thus, conditions placed on the tail probabilities are embodied in the domain, and a little linear algebra gives us our result, without additional domination-like conditions.

An advantage of our approach is that it makes immediately clear the relevant topology on the space of admissible probability measures. This is important because it makes clear what a perturbation of a probability measure looks like (though Ledyard, 1971 also presents a pseudo-metric on the space of probability measures), while simultaneously allowing for a perturbation of the utility function in a large class (this is not immediately obvious in Ledyard). Such perturbations are invaluable to analysts, and our approach lays bare the connection between the weight function, the spaces of admissible probability measures and vN–M utility functions, and the corresponding topologies on these spaces.

References

Blackwell, David A., Girshick, M.A., 1954. Theory of Games and Statistical Decisions. Wiley, New York. Dover Reprint 1979.

DeGroot, Morris, 1970. Optimal Statistical Decisions. McGraw-Hill, New York.

Fishburn, Peter C., 1975. Unbounded expected utility. Ann. Statist. 3 (4), 884–896.

Fishburn, Peter C., 1976. Unbounded utility functions in expected utility theory. Quart. J. Econ. 90 (1), 163–168.

Föllmer, Hans, Schied, Alexander, 2011. Stochastic Finance: An Introduction in Discrete Time. Walter de Gruyter, Berlin.

Grandmont, Jean-Michel, 1972. Continuity properties of a von Neumann–Morgenstern utility. J. Econom. Theory 4, 45–57.

Herstein, Israel Nathan, Milnor, John, 1953. An axiomatic approach to measurable utility. Econometrica 21 (2), 291–297.

Huang, Chi-fu, Litzenberger, Robert H., 1988. Foundations for Financial Economics. Prentice Hall, Upper Saddle River, NJ.

Kreps, David M., 1988. Notes on the Theory of Choice. Westview Press, Boulder, Co.

Kreps, David M., 2012. Microeconomic Foundations I: Choice and Competitive Markets. Princeton University Press, Princeton, NJ.

Ledyard, John O., 1971. A pesudo-metric space of probability measures and the existence of measurable utility. Ann. Math. Stat. 42, 794–798.

von Neumann, John, Morgenstern, Oskar, 1944. Theory of Games and Economic Behavior, first ed. Princeton University Press, Princeton, NJ.