The Pearson System of Utility Functions[*]

Marco LiCalzi
<licalzi@unive.it>
Dept. of Applied Mathematics
University of Venice

Annamaria Sorato
<amsorato@unive.it>
Dept. of Applied Mathematics
University of Venice

(November 2003)

Abstract. This paper describes a parametric family of utility functions for decision analysis. The parameterization is obtained by embedding the HARA class in a four-parameter representation for the risk aversion function. The resulting utility functions have only four shapes: concave, convex, S-shaped, and reverse S-shaped. This makes the family suited for both expected utility and prospect theory. We also describe an alternative technique to estimate the four parameters from elicited utilities, which is simpler and easier to implement than standard fitting by minimization of the mean quadratic error.

Keywords: coefficient of risk aversion, elicitation of preferences under risk, expected utility, HARA utility functions, Pearson system of distributions, prospect theory, probability weighting function, target-based decisions.

JEL Classification Numbers: D81.

Correspondence to:

Marco LiCalzi	Department of Applied Mathematics, University of Venice
	Dorsoduro 3825/e
	30123 Venezia, Italy
Phone:	[++39] (041)-234-6925
Fax:	[++39] (041)-522-1756
E-mail:	licalzi@unive.it

---

1 Introduction

Utility functions over monetary outcomes are essential to decision analysis, because they are often a necessary condition to guide or describe choices among risky gambles. In applications, the task of working with utility functions is usually made more manageable by restricting attention to a few convenient parametric families.

Under the expected utility paradigm, the dominant approach to constrain the set of admissible parameterizations goes through the assumption of risk aversion and leads to concave utility functions. The best known example is constant absolute (or relative) risk aversion, which limits the choice to linear and exponential (or power and logarithmic) functions. Other examples include proper risk aversion (Pratt and Zeckhauser, 1987), standard risk aversion (Kimball, 1993), and constant risk exchange (Farquhar and Nakamura, 1987). In a well-received proposal, Bell (1988) shows how the *one-switch rule* — increasing wealth cannot switch preferences among two risky gambles more than once — implies four functional forms for the utility function. Bell and Fishburn (2000) carries over and generalizes this argument to three non-expected utility models.

This approach does not work for the parameterization of the utility functions in prospect theory. An obstacle stems from the differential treatment of the assumption of *asset integration*. Expected utility maintains that the argument of the *utility function $u(w)$* is total wealth. Prospect theory contends that the "carriers of value are changes in wealth" and that the value is a function of two arguments: "the asset that serves as reference point, and the magnitude of the change"; see Kahneman and Tversky (1979, p. 277). Inspired by Markowitz (1952), they advocate a *value function $v(x; m)$* where x denotes increments in wealth and m is the reference point. In practice, prospect theory holds that preferences are not much affected by "small or even moderate variations" in the reference point and hence assumes that the value function depends only on the wealth increments x, writing it as $v(x)$.

The fundamental difficulty, of course, is that expected utility and prospect theory make conflicting assumptions about the risk attitude of an agent. According to prospect theory, the failure of asset integration engenders a *reflection effect* and *loss aversion*. The first one yields risk aversion over gains and risk propensity over losses. The second one induces a steeper slope of the utility function over losses. The two effects imply that $v(x)$ is S-shaped with a steeper slope below the reference point. A third problem is that a complete parameterization of prospect theory requires to deal also with the probability weighting function, increasing the difficulty of calibrating a model.

It is not surprising, therefore, that the existing parameterizations of prospect theory are not satisfactory. Neilson and Stowe (2002) makes a compelling argument that the proposed parameterizations may "fit the experimental data well, [but] they have poor out-of-sample performance" and concludes that "alternative functional forms are needed" (p. 44). The purpose of this paper is to describe a parameterization for utility and value functions that works across expected utility and prospect theory. In a nutshell, we work out a direct parameterization of the risk aversion functions — turn to Equation (5) below for a preview — and then recover from this the associated utility (or value) functions. As a byproduct of this technique, all the functions in our family are twice differentiable.

1

The most important advantages of the proposed parameterization are the following. First, it embeds all the HARA utility functions, which include the most commonly used functional forms in expected utility. Second, there are only four possible shapes for the functional forms: convex, concave, S-shaped and reverse S-shaped. Third, the parameterization requires at most four parameters; moreover, one of these can be directly interpreted as the reference point assumed by prospect theory. Fourth, the parametric family is the analog of a well-known parameterization for probability distribution functions; in particular, this suggests an estimation procedure which is easier to implement than the mean residual error technique typically used for utility assessment and calibration.

On the downside, our parameterization includes utility (or value) functions which have no closed-form representation. However, since their first derivatives always admit an explicit representation, they can be easily recovered by numerical integration. In a world where decision analysis is usually carried out by means of interactive computer software, we see this as an acceptable shortcoming. We also note that preference functionals like expected utility can still be analytically computed by means of a trivial integration by parts. The paper also discusses an alternative interpretation of expected utility that makes working with first derivatives very natural.

Finally, a word of caution: the arguments we provide to support the mathematical interest of the proposed parametric family have a theoretical bias. The acid test for a parameterization must come from empirical work comparing the performance of alternative families over tasks such as utility assessment and calibration. Since we lack a proper training as experimentalists, these important tests are better left where the comparative advantage lies.

The plan of the paper is the following. Section 2 presents a parameterization of the risk aversion function which is closely linked to the Pearson system of probability distributions. Section 3 provides a general description of the utility functions associated with this parameterization. Section 4 discusses the target-based interpretation of expected utility as a tool for the construction of preferences and demonstrates how to use simple statistics associated with the utility function to derive tests or describe psychological properties. Section 5 describes an estimation procedure for the parameters, which is known in the literature as the method of moments. Section 6 offers some concluding remarks, including a suggestion to extend the parameterization to probability weighting functions.

2   The Pearson equation

We begin with expected utility. Consider a von Neumann-Morgenstern utility function $u(w)$ defined over monetary outcomes, that we interpret as final wealth positions. The *support* of $u(w)$ is the (possibly unbounded) interval (w_1, w_2) of its points of increase. We assume that over its support $u(w)$ is bounded, strictly increasing and twice differentiable, with $u'(w) > 0$. Since $u(w)$ is unique up to a strictly increasing affine transformation (that is, it is defined on an interval scale), we normalize its range so that $\inf u(w) = 0$ and $\sup u(w) = 1$. When necessary, we extend the domain of u by continuity, taking $u(w) = 0$ for $w \leq w_1$ and $u(w) = 1$ for $w \geq w_2$.

Under our assumptions, the (local and absolute) *risk aversion function*

$$r(w) = -\frac{u''(w)}{u'(w)}$$

is properly defined over the same support as u. The interpretation of $r(w)$ is well-known. It equals twice the risk premium per unit of variance for an infinitesimal zero-mean risk at a given level of wealth w. The risk aversion function was first studied in deFinetti (1952) and independently rediscovered by Arrow (1971) and Pratt (1964).

Even if its definition is local, the risk aversion function characterizes preferences in a global sense as well. As Pratt (1964, p. 126) puts it, "the local risk aversion function r associated with any utility function u contains all essential information about u". More precisely, we can recover $u(w)$ from $r(w)$. Integrating $-r(w)$ and exponentiating gives $e^c u'(w)$. This can be integrated again to derive $e^c u(w) + d$, where c and d are two immaterial constants of integration. Consistent with our normalization, we choose the version of u' such that

$$\int_{w_1}^{w_2} u'(w)\,\mathrm{d}w = 1 \tag{1}$$

so that $c = d = 0$.

The one-to-one relationship between risk aversion functions and (normalized) utility functions makes it possible to advance a direct parameterization of the risk aversion function and derive from this a class of utility functions. The best known example for this approach is the class of HARA utility functions. They inherit this name from being associated with an hyperbolic risk aversion function, which (up to irrelevant constants) can be written as

$$r(w) = \frac{1}{\alpha + \beta w} \tag{2}$$

The HARA utility functions include as special cases the utility functions with constant absolute risk aversion for $\beta = 0$ (in particular, risk neutrality obtains for $\alpha \to \infty$) and those with constant relative risk aversion for $\alpha = 0$. They are often invoked as sufficient (and sometimes) necessary conditions in several theoretical results, ranging from linear risk sharing rules (Cass and Stiglitz, 1970) to the optimality of myopic policies in dynamic portfolio selection (Merton, 1969). Therefore, we require that our parametric family should embed the HARA functions as a special case.

Let us now turn to prospect theory. We assume that the support of $v(x)$ is a (possibly unbounded) interval over which $v(x)$ is bounded, strictly increasing and twice differentiable, with $v'(x) > 0$. Consistent with common practice and the axiomatization in Wakker and Tversky (1993), we also assume that $v(0) = 0$ and that $v(x)$ is defined on a ratio scale.

Since expected utility deals with functions over total wealth and prospect theory deals with functions over wealth increments, we need to bridge the gap between utility functions and value functions. This requires to specify how to carry out asset integration. We assume that the effects of the reference point and of the wealth increments are additively separable. Hence, given the value function $v(x)$ over wealth increments and a reference level m, the utility function $u(w)$ over the total wealth $w = m + x$ can be recovered as

$$u(m + x) = u(m) + v(x). \tag{3}$$

Note that, although this assumption implies that a lottery maximizes the expected value of $u(m + x)$ if and only if it maximizes the expected value of $v(x)$, the two theories may not generate the same preferences under risk. Prospect theory allows for probability distortions and therefore we cannot presume that the expected value is computed using the stated probabilities. Consequently, we do not claim equivalence of choice behavior under the two theories.

Define $\hat{r}(x) = -v''(x)/v'(x)$ as a formal equivalent of the risk aversion function for $v(x)$. The function $\hat{r}(x)$ cannot be interpreted as a risk aversion function, because probability distortions may occur under prospect theory. However, after our normalization, Equation (3) implies $u'(m + x) = v'(x)$. This has two important consequences. First, the recovery of $e^c v(x) + d$ from $\hat{r}(x)$ is analogous; in particular, we choose the version of v' such that

$$\int_{w_1 - m}^{w_2 - m} v'(x)\, \mathrm{d}x = 1$$

so that $c = d = 0$.

Second, we can unify the treatment of the two paradigms by dealing with the function $r(m + x)$, where m is a reference level and $w = m + x$ is the final wealth position. Within the expected utility paradigm, the function $r(m + x) = r(w)$ is interpreted as a risk aversion function from which we can recover the utility function $u(w)$. Within the prospect theory paradigm, $r(m + x) = \hat{r}(x)$ is no more than the formal equivalent of a risk aversion function, from which we can identify the utility function $u(m + x)$ by carrying out asset integration under (3). For convenience, regardless of the paradigm, from now on we call r a risk aversion function.

Under prospect theory, the most representative property of the value function is its S-shape, which is usually justified as a consequence of diminishing sensitivity. Since $v(x)$ is concave if and only if $r(m + x)$ is positive, a parameterization for r consistent with an S-shape for v must allow for a case in which the sign of $r(m + x)$ is equal to the sign of x. We assume linearity and require that our parameterization embeds the equation

$$r(m + x) = \frac{x}{\gamma} \tag{4}$$

Note that $\gamma \to +\infty$ corresponds to the linear case. Moreover, $\gamma < 0$ would generate a reverse S-shape. In light of the controversial case raised in its favor by Levy and Levy (2002), we prefer not to impose $\gamma > 0$.

Equations (2) and (4) represent the two special cases that we want to embed in our parameterization of the risk aversion function. They correspond respectively to an hyperbolic and to a linear functional form. Taking reciprocals, these would respectively be transformed into a linear and a hyperbolic function. Therefore, it only seems natural to attempt combining them through an arithmetic or an harmonic mean. We choose the latter because it yields a rich parameterization that is also widely studied in statistics. After renaming constants, the harmonic mean of $r(w) = 1/(\alpha + \beta w)$ from (2) and $r(w) = (w - m)/\gamma$ from (4) can be written as

$$r(w) = \frac{w - m}{aw^2 + bw + c}. \tag{5}$$

4

This is the parametric family of risk aversion functions discussed in this paper. Section 3 presents the main utility functions associated with this family. (From now, we subsume the value functions under the heading of utility functions for simplicity.) We discuss here the relationship of (5) with a well-known parameterization used in statistics.

Between 1890 and 1895, Pearson (1895; see also 1901, 1916) worked out a system of probability distributions in which every member has a density function $f(x)$ which satisfies the differential equation

$$\frac{\mathrm{d} \log f(x)}{\mathrm{d}x} = -\frac{x - m}{ax^2 + bx + c}. \tag{6}$$

His purpose was to offer a four-parameter family of distributions from which one could select a member to approximate empirical distributions. The similarity with the problem of designing a parametric family for utility assessment and calibration runs deeper than one might suspect by comparing (5) and (6).

Under our assumptions, $u(w)$ is a bounded, strictly increasing and twice differentiable function defined over $\mathbb{R}$ and normalized between 0 and 1. Moreover, $u'(w)$ integrates to 1. Hence, $u(w)$ is formally equivalent to a probability distribution function and $u'(w)$ to a probability density function. This observation recurs in the literature. Borch (1968) uses it to study the probability of ruin. Berhold (1973) exploits it to propose a family of natural-conjugate utility functions inspired by results in Bayesian statistics. Castagnoli and LiCalzi (1996) shows that this formal equivalence can be given meaning by interpreting expected utility as the probability of meeting a target which is imperfectly known. Bordley and LiCalzi (2000) builds on this to propose a target-based paradigm which preserves the results of the expected utility model by giving up the notion of a utility function; see Section 4, or Bordley (2002) for a more extensive discussion.

When $u(w)$ is interpreted as a probability distribution function, $u'(w)$ turns out to be its density function. Since $r(w) = -\mathrm{d} \log u'(w)/\mathrm{d}w$, (5) and (6) are formally equivalent. For this reason, we refer to them respectively as the *Pearson equation* for risk aversion functions and the Pearson equation for densities. The estimation technique for assessing utility functions suggested in this paper is the adaptation of Pearson's *method of moments* for the selection of distribution functions. This technique is described in Section 5.

3   The Pearson system of utility functions

The solutions to (6) are often nicknamed *Pearson densities*. By analogy, we call *Pearson utility functions* those associated with the risk aversion functions described by (5). The purpose of this section is to provide an appreciation of the main properties and of the diversity of the set $\mathcal{P}$ of Pearson utility functions, which includes as special cases almost all the most important examples commonly used in the literature — the biggest exception being some of the one-switch functions proposed in Bell (1988).

A convenient way to approach the study of the Pearson system of utility functions is to rewrite (5) as

$$u''(w) = -\left[\frac{(w - m)}{aw^2 + bw + c}\right] u'(w) \tag{7}$$

and to work out the properties of u by studying the solutions u' of (7). Note that we are only interested in those solutions where $u'(w) \geq 0$ over the support of u — corresponding to an increasing utility function. Hence, formal solutions of (7) may be restricted to an interval (w_1, w_2) where $u'(w) \geq 0$, assigning the value $u'(w) = 0$ outside of the interval.

A few properties follow immediately. First, since $u'(w)$ cannot be strictly negative, a Pearson utility function can have at most one inflexion point at m. This implies that the risk attitude can switch at most once. Hence, a Pearson utility function has only one of four shapes: concave, convex, S-shaped and reverse S-shaped. These are exactly the typical cases considered in most of the literature. Note also that the risk attitude switches if and only if the reference point m belongs to the interval (w_1, w_2). Therefore, any technique which selects a "best fit" among the utility functions associated with (7) provides a simple method for deriving m endogenously. Contrast this with prospect theory, which bypasses the issue of estimating the reference point by assuming $m = 0$; this choice reduces the number of free parameters in (7) to three.

Second, if u is a Pearson utility function over an interval S, it is also a Pearson utility function over any subinterval $S' \subset S$. To avoid ambiguities, we always assume the largest possible support. However, in applications where only a few values for the utility function have been elicited, this property allows one to estimate or calibrate a Pearson utility function over a smaller interval and then consistently extend it — if necessary — to a larger support. As we see below, the maximal support for a Pearson utility function may range from a bounded interval to $\mathbb{R}$, depending on the combination of the parameters.

A third useful property is that the set $\mathcal{P}$ of Pearson utility functions is invariant to shift and rescaling of the x-units: if $u_1(x)$ belong to $\mathcal{P}$, so does $u_2(x) = u_1(\alpha x + \beta)$ with $\alpha > 0$. This condition is the equivalent for utility functions of the location-scale conditions for random variables in Meyer (1987). Moreover, $\mathcal{P}$ is also invariant under an order reversal: that is, if $u_1(x)$ is in $\mathcal{P}$, so does $u_2(x) = u_1(\alpha x + \beta)$ for any $\alpha \neq 0$. This last property makes it possible to describe P by partitioning it in equivalence classes, each of which is closed under shifts, rescaling and order reversal.

Pearson himself organized the solution to his equation in a system of twelve classes identified by a number. His numbering criterion has no systematic basis, but it has nonetheless become customary to use it and thus we retain it. Since some of the original Pearson's classes can be retrieved as limiting cases of other classes, we follow Johnson et alii (1994) and explicitly discuss only the major cases. See Kendall (1948) for a more detailed exposition. To circumvent the nonexistence of closed-form solutions for some classes of utility functions, we list $u'(w)$ as representative element. Recall the interpretation of u' as the analog of the density function of a distribution function. When possible, we use this analogy to suggest a nickname for each class.

The form of the solution to (7) depends on the nature of the zeros of the denominator $(aw^2 + bw + c)$ on the right-hand side. In the special case $a = b = 0$, we obtain

$$u'(w) = K \exp \left[-\frac{(w-m)^2}{2c} \right] \qquad \text{for } x \in \mathbb{R}, \tag{8}$$

where $K = (\sqrt{2\pi c})^{-1}$ is a constant chosen to normalize the integral of $u'(w)$ to 1. The exact value of $K > 0$ is irrelevant for decision analytic purposes, so we omit its specification

in the following. The resulting *normal* utility function is not assigned to any class and can be viewed as a limiting case. Except for III, assume $a \neq 0$ in the following discussion.

I. (Beta) Suppose $\Delta := b^2 - 4ac > 0$ and assume that the two real roots $r_1 < r_2$ of $(aw^2 + bw + c) = 0$ lie on opposite sides of the support of u. Then the representative element is

$$u'(w) = K(w - r_1)^{k_1}(r_2 - w)^{k_2} \qquad \text{for } w \in (r_1, r_2),$$

with $k_1 = (r_1 - m)/[a(r_2 - r_1)] > -1$ and $k_2 = (m - r_2)/[a(r_2 - r_1)] > -1$. (We leave it understood that $u'(w) = 0$ elsewhere.)

The shape of the corresponding utility function u depends on the signs of k_1 and k_2: if $k_1 \cdot k_2 < 0$, it is concave when $k_1 < 0$ and convex when $k_1 > 0$; if $k_1 \cdot k_2 > 0$, it is S-shaped when $k_1 > 0$ and reverse S-shaped when $k_1 < 0$. The symmetric case with $k_1 = k_2 \neq 0$ is known as Class II. The risk-neutral utility function $u(w) = w$ is the limiting case associated with $k_1 = k_2 \to 0$.

III. (Gamma) Suppose $a = 0$ and $b \neq 0$. For $c > 0$, the representative element is

$$u'(w) = K(c + bw)^k \exp\left(-\frac{w}{b}\right) \qquad \text{for } w > -\frac{c}{b},$$

with $k = (c/b^2) + (m/b)$. The corresponding utility functions are S-shaped if $k > 0$ and concave if $k < 0$. For $c < 0$, the support becomes $w < -(c/b)$ and the corresponding utility functions are reverse S-shaped for $k > 0$ and convex for $k < 0$. The case $k = 0$ gives Class X, which is formed by the exponential utility functions.

IV. Suppose $\Delta < 0$. Then the representative element is

$$u'(w) = K\left[C + a(w + B)^2\right]^{-\frac{1}{2a}} \exp\left[\frac{m + B}{\sqrt{aC}}\arctan\frac{w + B}{\sqrt{(C/a)}}\right] \qquad \text{for } w \in \mathbb{R},$$

where $B = b/(2a)$ and $C = c - [b^2/(4a)] = c - B^2$. The corresponding utility functions are S-shaped.

V. (Inverse gamma) Suppose $\Delta = 0$ and let r be the only root of $(aw^2 + bw + c) = 0$. When $(m - r)/a > 0$, the representative element is

$$u'(w) = K(w - r)^{-\frac{1}{a}} \exp\left(-\frac{m - r}{a(w - r)}\right) \qquad \text{for } w > r.$$

The corresponding utility functions are S-shaped for $a > 0$ and concave for $a < 0$. When $(m - r)/a < 0$, the representation is similar but the support is $w < r$. Moreover, the corresponding utility functions are reverse S-shaped for $a > 0$ and convex for $a < 0$. Utility functions with constant relative risk aversion belong to this class for $m = r = 0$. Moreover, when $m = r$, and $|a| < 1$, we have the special case $u'(w) = K(w - r)^{-1/a}$. This is respectively known as Class VIII and Class IX for $a > 0$ and $a < 0$.

VI. (Inverse beta) Suppose $\Delta > 0$ and assume that the two real roots $r_1 < r_2$ of $(aw^2 + bw + c) = 0$ lie on the same side of the support of u. Assuming a support on the right of r_2, the representative element is

$$u'(w) = K(w - r_1)^{k_1}(w - r_2)^{k_2} \qquad \text{for } w > r_2,$$

with $k_2 < -1$ and $k_1 + k_2 < 0$. The corresponding utility functions are concave. A similar representation holds when the support lies on the left of r_1, but in this case the utility functions are convex. This class bears a relationship with Class I analogous to the one between III and V.

VII. Suppose $m = b = 0$, $a > 0$ and $c > 0$. Then the representative element is

$$u'(w) = K(c + aw^2)^{-(1/2a)} \qquad \text{for } w \in \mathbb{R}.$$

The corresponding utility functions are S-shaped.

The representative element for each of these classes is the derivative of the utility function u. This looks unusual, and possibly hard to interpret. We have two comments in this respect. The first one is that there is no loss of generality (except for the assumption of differentiability itself) or computational ability. A simple integration by parts shows that knowledge of the functional form of u' can replace knowledge of the functional form of u in the computation of the expected utility of a risky lottery X. (The proof is in the appendix.)

Proposition 1 *Let u be differentiable. Given a random variable X with distribution function F, suppose that $E|u(X)| < \infty$. Then*

$$Eu(X) = u(0) - \int_{-\infty}^{0} u'(x)F(x)\,dx + \int_{0}^{+\infty} u'(x)[1 - F(x)]\,dx.$$

A similar argument applies to several (even if not all) the different versions of prospect theory. For instance, consider the cumulative prospect theory of Tversky and Kahneman (1992) when the two nonadditive measures w^+ and w^- over gains and losses can be written as two (possibly different) distortions $g^+(\cdot)$ and $g^-(\cdot)$ of the underlying probability measure. In this case (recall that $v(0) = 0$), the preference functional is

$$\int_{-\infty}^{0} v(x)\,d(g^- \circ F)(x) - \int_{0}^{+\infty} v(x)\,d[g^+ \circ (1 - F)](x)$$

which, upon integration by parts, becomes

$$-\int_{-\infty}^{0} v'(x)(g^- \circ F)(x)\,dx + \int_{0}^{+\infty} v'(x)[g^+ \circ (1 - F(x))]\,dx.$$

The second — and, in our eyes, more important — comment is that there exists a plausible point of view which actually makes working with u' the natural option for assessing and calibrating a preference functional under risk. We turn to illustrate it in the next section.

4   Targets and simple statistics

Let us start with an example. Suppose that the agent is trying to decide which of several risky investments he should commit to in view of saving for his retirement. The agent does not know which investment he likes best and therefore must think his way about the best choice. Here is one possible route. Suppose that the agent can pinpoint the target wealth t that he would like to achieve by the time of his retirement. In this ideal situation, he should pick an investment which delivers a final wealth X that maximizes the probability $P(X \geq t)$ of meeting his target.

However, at the moment of making his choice, the agent generally lacks crucial information about the exact amount of the target he should seek. In other words, the target is not perfectly known. We can model this by means of a random variable $T = t + \varepsilon$, where ε is a zero-mean error term that is stochastically independent of the risky options available to the agent. The natural extension of the previous criterion is to pick an investment that, at the moment of making the choice, maximizes the probability $P(X \geq T)$ of meeting the imperfectly known target.

Suppose that the (stochastically independent) target T has distribution function $u(x)$ and that the final wealth X associated with an investment has distribution function $F(x)$. Then

$$ P(X \geq T) = \int P(x \geq T) \, \mathrm{d}F(x) = \int u(x) \, \mathrm{d}F(x). $$

Since the mathematical representation of the functional is the same, an agent trying to maximize the probability of meeting a target with distribution function $u(x)$ would make exactly the same choices as another agent trying to maximize the expected value of a utility function u. In other words, target-based decisions and expected utility decisions are observationally equivalent. Therefore, these two competing viewpoints have the same explanatory power, and it is a matter of taste (or convenience) which one to adopt.

Under the target-based perspective, the function $u(w)$ is the cumulative distribution function of an imperfectly known target. Therefore, working with $u'(w)$ is equivalent to dealing with its probability density function. Entrenched usage favors a value-based language and therefore suggests to assess a utility function u. The target-based perspective brings to the fore a probability-based language and ends up dealing with the density function u' of the target. There are advantages in exploring this alternative point of view. See also Bordley and LiCalzi (2000).

For instance, consider the reflection effect of prospect theory. Arguing from a value-based perspective, it is defined by the requirement that the value function $v(x)$ is S-shaped. This mathematical characterization is sufficient if prospect theory has only the modest goal of accounting for what people actually do. However, a more compelling approach should also explain how people construct their preferences whenever they do not happen to know them already. In this respect, the target-based viewpoint suggests that, if people are trying to "beat" some implicitly defined target, the reflection effect is the consequence of estimating a density function $u'(m + x)$ which is unimodal around the reference point m.

Given the scope of this paper, this is not the place for an exhaustive comparison of the two viewpoints. We have chosen to couch the presentation in the language of utility,

because this is the established convention in the profession. The short discussion above in this section is meant to suggest that an alternative and perhaps more natural route would be to turn the problem of providing a parameterization for u into the problem of parameterizing the density u' of an imperfectly known target.[1]

There are at least two other advantages in viewing u as a probability distribution function. First, this may help to devise new experimental tests. The key point is to compute and make use of simple statistics such as the mode or the median or the moments associated with u. Note that working with these statistics is (consistent with but) independent of the target-based interpretation. For instance, while the target-based interpretation provides an argument to define the reference point m as the mode of the imperfectly known target, the definition may stand on its own. For simplicity, we equate the reference point and the mode associated with u in the following example.

An important tenet of prospect theory is the notion of loss aversion, which substantiates the idea that losses matter more than gains: an agent usually turns down fair gambles around his reference point. The common rendition formalizes this property by the assumption that $u'(m + x) \leq u'(m - x)$ for $x > 0$. However, as Kahneman and Tversky (1979) is careful to show, a slightly weaker assumption suffice:

$$u(m) - u(m - x) \geq u(m + x) - u(m) \qquad \text{for all } x > 0. \tag{9}$$

In words, the increment in utility from reducing a loss by x is greater than (or equal to) the increment in utility from making a gain of equal amount. (See Neilson (2002) for an alternative and stronger definition of loss aversion.)

If we think of u as a distribution function and let $x \to +\infty$ in (9), we obtain $u(m) \geq 1 - u(m)$; that is, $u(m) \geq 1/2$. Therefore, a necessary condition for u to exhibit loss aversion is that the median associated with u should not be greater than its mode. This is a simple and general property that can be tested using elicited utility functions. Under the additional assumption that u is a Pearson utility function, we can derive a second test. Sato (1997) has shown that unimodal Pearson densities with finite mean always satisfy the mean-median-mode inequality, which states that the median is always contained in the (possibly, degenerate) interval delimited by the mean and the mode. Therefore, a necessary condition for S-shaped utility functions to exhibit loss aversion is that their (finite) first moment should not be greater than the mode.

The second type of advantage is that we can use the statistics associated with u to provide simple but effective indicators for psychological patterns regarding preferences under risk. Here is an example. Consider the normal utility function given in Equation (8). The two parameters m and c describe respectively the position of the reference point and the intensity of the diminishing sensitivity effect. The greater m, the larger the reference point; analogously, the smaller c, the sharper the changes in utility associated with increments in wealth from the reference point. On the other hand, m and c are easily recognized as the mean and the variance associated with u. Hence, at least for this type of utility functions, the second (central) moment can be used to measure the strength of the diminishing sensitivity

---

[1] A bit of reverse engineering also shows how much this paper has borrowed from Pearson's approach about fitting theoretical density functions to empirical distributions.

effect. The next section builds on the moments associated with u to derive a procedure to fit Pearson utility functions to elicited utilities.

5   The method of moments

In general, the assessment of utility functions over money is carried out in three stages: first, one elicits a few values for the utility function; second, one selects a parametric class of utility functions; third, one estimates the parameters by searching for the best fit with the data. This section presents a method to handle the third stage for the Pearson family of utility functions. As we discuss below, this *method of moments* can be used in alternative or in conjunction with the standard approach that estimates parameters by minimizing the sum of squared errors or some analogous measure.

As for the first stage, under the expected utility paradigm there exist several methods for the elicitation of utilities; see Farquhar (1984) for a survey. However, the most reliable among these techniques require the assumption that the agent does not distort the probabilities of the risky outcomes. This assumption is untenable, as the successful challenge of prospect theory to the descriptive validity of expected utility has made clear. In a laudable effort to make utility elicitation independent of the paradigm, Wakker and Deneffe (1996) have thus developed the *tradeoff method*, which is not affected by probability distortions. This advantage has been promptly recognized and endorsed by the advocates of prospect theory; see for instance Abdellaoui (2000), Bleichrodt and Pinto (2000), and Fennema and van Assen (1999). Hence, utility elicitation is by now essentially unified across the paradigms of expected utility and prospect theory.

In general, the method of moments is a statistical technique to estimate probability distributions by equating their theoretical moments with the moments of the empirical distributions. When applied to the estimate of Pearson distributions, it consists in choosing the four parameters in (6) to make sure that the theoretical moments of the associated distribution function match the empirical moments. We present the method with reference to the Pearson equation (6) for the density function and then show how to apply it to the estimate of the coefficients of the Pearson equation (5) for the risk aversion function.

Given a density function $f(x)$ that satisfies the Pearson equation (6), suppose that the first four moments $\mu_k = \int x^k f(x)\,\mathrm{d}x$ for $k = 1, \ldots, 4$ exist and are finite. Using (6), a bit of manipulation — see Johnson et alii (1994, p. 22) — leads to the system of linear equations

$$\begin{array}{ccccccccc}
m & + & 2\mu_1 a & + & b & & & = & \mu_1 \\
\mu_1 m & + & 3\mu_2 a & + & 2\mu_1 b & + & c & = & \mu_2 \\
\mu_2 m & + & 4\mu_3 a & + & 3\mu_2 b & + & 2\mu_1 c & = & \mu_3 \\
\mu_3 m & + & 5\mu_4 a & + & 4\mu_3 b & + & 3\mu_2 c & = & \mu_4
\end{array}, \tag{10}$$

from which we obtain the four parameters m, a, b, and c as a function of the first four

moments:

$$m = \frac{9\mu_1\mu_2^3 + 8\mu_1\mu_3^2 + 12\mu_1^3\mu_4 + 3\mu_2^2\mu_3 + \mu_3\mu_4 - 20\mu_1^2\mu_2\mu_3 - 13\mu_1\mu_2\mu_4}{18\mu_2^3 + 12\mu_3^2 - 6\mu_1^2\mu_2^2 + 8\mu_1^3\mu_3 + 10\mu_1^2\mu_4 - 10\mu_2\mu_4 - 32\mu_1\mu_2\mu_3}$$

$$a = \frac{6\mu_2^3 + 3\mu_3^2 - 3\mu_1^2\mu_2^2 + 4\mu_1^3\mu_3 + 2\mu_1^2\mu_4 - 2\mu_2\mu_4 - 10\mu_1\mu_2\mu_3}{18\mu_2^3 + 12\mu_3^2 - 6\mu_1^2\mu_2^2 + 8\mu_1^3\mu_3 + 10\mu_1^2\mu_4 - 10\mu_2\mu_4 - 32\mu_1\mu_2\mu_3}$$

$$b = \frac{-3\mu_1\mu_2^3 - 2\mu_1\mu_3^2 - 6\mu_1^3\mu_4 - 3\mu_2^2\mu_3 - \mu_3\mu_4 + 8\mu_1^2\mu_2\mu_3 + 7\mu_1\mu_2\mu_4}{18\mu_2^3 + 12\mu_3^2 - 6\mu_1^2\mu_2^2 + 8\mu_1^3\mu_3 + 10\mu_1^2\mu_4 - 10\mu_2\mu_4 - 32\mu_1\mu_2\mu_3}$$

$$c = \frac{-4\mu_1^2\mu_3^2 + 3\mu_2\mu_3^2 - 4\mu_2^2\mu_4 + 3\mu_1^2\mu_2\mu_4 + \mu_1\mu_2^2\mu_3 + \mu_1\mu_3\mu_4}{18\mu_2^3 + 12\mu_3^2 - 6\mu_1^2\mu_2^2 + 8\mu_1^3\mu_3 + 10\mu_1^2\mu_4 - 10\mu_2\mu_4 - 32\mu_1\mu_2\mu_3}$$

$$(11)$$

We warn that statisticians are used to assume $\mu_1 = 0$ and solve for the parameters as a function of the central moments or other statistics. Therefore, formulas taken from handbooks in statistics will in general look different (and probably nicer). Here is an example, with μ_1 possibly different from 0. Let X be a random variable with density $f(x)$. Let μ and σ denote the mean and the standard deviation of X and denote the two coefficients of skewness and kurtosis for X by

$$\beta_1 = \frac{E(X - \mu)^3}{\sigma^3} \quad \text{and} \quad \beta_2 = \frac{E(X - \mu)^4}{\sigma^4} - 3.$$

Then the solution in (11) can be rewritten as

$$m = \mu + \frac{\sigma\beta_1(\beta_2 + 6)}{12\beta_1^2 - 10\beta_2 - 12}$$

$$a = \frac{3\beta_1^2 - 2\beta_2}{12\beta_1^2 - 10\beta_2 - 12}$$

$$b = -\frac{\sigma\beta_1(\beta_2 + 6) + 2\mu(3\beta_1^2 - 2\beta_2)}{12\beta_1^2 - 10\beta_2 - 12}$$

$$(12)$$

$$c = \frac{\mu^2(3\beta_1^2 - 2\beta_2) + \mu\sigma\beta_1(\beta_2 + 6) + \sigma^2(3\beta_1^2 - 4\beta_2 - 12)}{12\beta_1^2 - 10\beta_2 - 12}$$

In particular, given m and a, the values of b and c can be obtained as

$$b = \mu(1 - 2a) - m$$

$$c = \mu_2 a + (m - \mu)\left[\mu - \frac{2\sigma}{\beta_1}\right]$$

$$(13)$$

We now turn to the estimate of the risk aversion function in (5). Let there be a set of elicited utilities; formally, for some $w_1 < w_2 < \ldots < w_n$, we are given the corresponding

values of the utility function $u(w_1) \leq u(w_2) \leq \ldots \leq u(w_n)$. To rule out the trivial case of a constant utility function, we assume that at least one strict inequality holds. Apply the standard normalization and rescale utilities so that $u(w_1) = 0$ and $u(w_n) = 1$. Alternatively, one may specify one or two additional points $w_0 < w_1$ and $w_{n+1} > w_n$ and let $u(w_0) = 0$ and $u(w_{n+1}) = 1$.

We need to compute the "moments" associated with this utility function. Recall that $u(w)$ shares the formal properties of a distribution function but at this stage we only know a few points on its curve. Thus, we approximate it by the piecewise linear function $\hat{u}(w)$ joining $(w_i, u(w_i))$ to $(w_{i+1}, u(w_{i+1}))$ for all $i = 1, 2, \ldots, n - 1$. Using the corresponding piecewise constant "density"

$$\hat{u}'(w) = \frac{\hat{u}(w_{i+1}) - \hat{u}(w_i)}{w_{i+1} - w_i} \qquad \text{for } w \text{ in } (w_i, w_{i+1}),$$

we obtain the four moments by computing

$$\mu_k = \int_{w_1}^{w_n} x^k \hat{u}'(w) \, \mathrm{d}w \qquad \text{for } i = 1, 2, 3, 4.$$

These moments are substituted in (11) — or, equivalently, in (12) — to derive the four parameters m, a, b, c and, possibly, other statistics such as σ or β_1. Once the values of the parameters have been obtained, a glance to the list of the major classes of Pearson utilities in Section 5 suffices to identify the functional form for $u'(w)$. In general, then, numerical integration is necessary to reconstruct u from u' because some of the Pearson utility functions have no closed-form representation.

As described above, the method of moments simultaneously selects a functional form and its exact parametric specification. This provides an elementary way to estimate the utility function. On the other hand, the method has little theoretical appeal except for its simplicity. (Since the Thirties, for instance, it has been replaced as the method of choice for the estimate of density functions in statistics by the maximum likelihood method suggested by Fisher.) An alternative possibility is to use it only to select a functional form from the classes in Section 5 and then apply standard techniques such as the minimization of mean square errors for the choice of the exact values of the parameters.

6 Closing remarks

Prospect theory (and its variants) presume the existence of a probability weighting function g. Recent formulations require that g is an increasing and continuous function that maps from $[0, 1]$ onto $[0, 1]$. For instance, the two best known parameterizations for g are $g(p) = (p^k)/[p^k + (1 - p)^k]^{1/k}$ in Tversky and Kahneman (1992) and $g(p) = \exp[-(-\ln p)^k]$ in Prelec (1998). Both of these functions are reverse S-shaped to fit the empirical evidence.

By analogy with the treatment of utility functions, one is naturally led to consider probability weighting functions that satisfy the Pearson equation. Given the additional constraint that the support of g is $[0, 1]$, this implies that the only possible choice belongs to Class I (Beta) and therefore suggests to look into a probability weighting function g

such that $g'(p) = [1/B(\alpha, \beta)]x^\alpha(1 - x)^\beta$. Adding the constraint of reverse S-shapedness imposes $\alpha < 0$ and $\beta < 0$. This derivation of a possible parameterization is purely formal, and at this stage we are unable to offer a probabilistic interpretation for $g(p)$ similar to the target-based interpretation of u. On the other hand, it adds plausibility to our modest claim that the Pearson equation is a natural source of alternative parameterizations.

Finally, we wish to point out that the limited number of parameters characterizing Pearson utility functions simplifies the identification of specific conditions over preferences. For instance, consider standard risk aversion by Kimball (1993). It is well known that a sufficient condition for a utility function u to exhibit standard risk aversion is that its coefficient of risk tolerance (i.e., the reciprocal of the risk aversion function) is increasing and concave — see Gollier (2001, p. 166). Imposing these two conditions on the reciprocal of the risk aversion function in (5), we obtain that $a > 0$ and $am^2 - bm - c \leq 0$ imply that the utility function $u(m + x)$ exhibits standard risk aversion over positive increments x above the reference point. In particular, if we assume $m = 0$ as in prospect theory, $a > 0$ and $c \leq 0$ suffice to ensure that the value function exhibits standard risk aversion over gains.

A   Proofs

PROOF OF PROPOSITION 1. By integration by parts,

$$\int_0^b u(x)\, dF(x) = u(0)[1 - F(0)] - u(b)[1 - F(b)] + \int_0^b u'(x)[1 - F(x)]\, dx \qquad (14)$$

Since $E|u(X)| < \infty$,

$$E|u(X)| = \sum_{k=0}^{+\infty} \int_k^{k+1} |u(x)|\, dF(x) < \infty$$

implies

$$\sum_{k=b}^{+\infty} \int_k^{k+1} |u(x)|\, dF(x) \to 0$$

as $b \uparrow \infty$. Thus, as $b \uparrow \infty$,

$$0 \leq u(b)[1 - F(b)] \leq |u(b)|P(|X| \geq b) \leq \sum_{k=b}^{+\infty} \int_k^{k+1} |u(x)|\, dF(x) \to 0$$

Going to the limit in (14),

$$\int_0^{+\infty} u(x)dF(x) = u(0)[1 - F(0)] + \int_0^{+\infty} u'(x)[1 - F(x)]\, dx$$

By a similar argument,

$$\int_{-\infty}^0 u(x)\, dF(x) = u(0)F(0) - \int_{-\infty}^0 u'(x)F(x)\, dx$$

from which the result follows. $\square$

References

[1] M. Abdellaoui (2000), "Parameter-free elicitation of utility and probability weighting function", *Management Science* **46**, 1497–1512.

[2] K.J. Arrow (1971), *Essays in the Theory of Risk Bearing*, Chicago: Markham Publishing.

[3] D. Bell (1988), "One-switch utility functions and a measure of risk", *Management Science* **34**, 1416–1424.

[4] D.E. Bell and P.C. Fishburn (2000), "Utility functions for wealth", *Journal of Risk and Uncertainty* **20**, 5–44.

[5] M.H. Berhold (1973), "The use of distribution functions to represent utility functions," *Management Science* **19**, 825–829.

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

[7] K. Borch (1968), "Decision rules depending on the probability of ruin," *Oxford Economic Papers* **20**, 1–10.

[8] R.F. Bordley (2002), "Foundations of target-based decision theory", in: H.W. Brachinger and P.A. Monney (eds.), *Decision Analysis*, from: Encyclopedia of Life Support Systems (EOLSS), Oxford: Eolss Publishers. [http://www.eolss.net]

[9] R.F. Bordley and M. LiCalzi (2000), "Decision analysis using targets instead of utility functions", *Decisions in Economics and Finance* **23**, 53–74.

[10] D. Cass and J. Stiglitz (1970), "The structure of investor preferences and asset returns, and separability in portfolio allocation", *Journal of Economic Theory* **2**, 122–160.

[11] E. Castagnoli e M. LiCalzi (1996), "Expected utility without utility", *Theory and Decision* **41**, 281–301.

[12] B. deFinetti (1952), "Sulla preferibilità", *Giornale degli Economisti e Annali di Economia* **11**, 685–709.

[13] P.H. Farquhar (1984), "Utility assessment methods", *Management Science* **30**, 1283-1300.

[14] P.H. Farquhar and Y. Nakamura (1987), "Constant exchange risk properties", *Operations Research* **35**, 206-214.

[15] H. Fennema and M. van Assen (1999), "Measuring the utility of losses by means of the tradeoff method", *Journal of Risk and Uncertainty* **17**, 277–295.

[16] C. Gollier (2001), *The Economics of Risk and Time*, Cambridge (MA): The MIT Press.

[17] N.L. Johnson, S. Kotz and N. Balakrishnan (1994), *Continuous Univariate Distributions, vol. I*, New York: Wiley, second edition.

[18] D. Kahneman and A. Tversky (1979), "Prospect theory: An analysis of decision under risk", *Econometrica* **47**, 263–291.

[19] M.G. Kendall (1948), *The Advanced Theory of Statistics*, London: Griffin, fourth edition.

[20] M. Kimball (1993), "Standard risk aversion", *Econometrica* **61**, 589-611.

[21] M. Levy and H. Levy (2002), "Prospect theory: Much ado about nothing?", *Management Science* **48**, 1334-1349.

[22] H. Markowitz (1952), "The utility of wealth", *Journal of Political Economy* **60**, 151–156.

[23] R.C. Merton (1969), "Lifetime portfolio selection under uncertainty: The continuous-time case", *Review of Economics and Statistics* **51**, 247–257.

[24] J. Meyer (1987), "Two-moment decision models and expected utility maximization", *American Economic Review* **77**, 421–430.

[25] W.S. Neilson (2002), "Comparative risk sensitivity with reference-dependent preferences", *Journal of Risk and Uncertainty* **24**, 131–142.

[26] W.S. Neilson and J. Stowe (2002), "A further examination of cumulative prospect theory parameterizations ", *Journal of Risk and Uncertainty* **24**, 31–46.

[27] K. Pearson (1895), "Memoir on skew variation in homogenous material", *Philosophical Transactions of the Royal Society A* **186**, 343–414.

K. Pearson (1901), "Supplement to a memoir on skew variation", *Philosophical Transactions of the Royal Society A* **197**, 443–459.

K. Pearson (1916), "Second supplement to a memoir on skew variation", *Philosophical Transactions of the Royal Society A* **216**, 429–457.

[28] D. Prelec (1998), "The probability weighting function", *Econometrica* **66**, 497–527.

[29] J.W. Pratt (1964), "Risk aversion in the small and in the large", *Econometrica* **32**, 122–136.

[30] J.W. Pratt and R.J. Zeckhauser (1987), "Proper risk aversion", *Econometrica* **55**, 143–154.

[31] M. Sato (1997), "Some remarks on the mean, median, mode and skewness", *Australian Journal of Statistics* **39**, 219–224.

[32] A. Tversky and D. Kahneman (1992), "Advances in prospect theory: Cumulative representation of uncertainty", *Journal of Risk and Uncertainty* **5**, 297–323.

[33] P. Wakker and D. Deneffe (1996), "Eliciting von Neumann-Morgenstern utilities when probabilities are distorted or unknown", *Management Science* **42**, 1131–1150.

[34] P. Wakker and A. Tversky (1993), "An axiomatization of cumulative prospect theory", *Journal of Risk and Uncertainty* **7**, 147–176.