University of Leicester

DEPARTMENT OF ECONOMICS

Probability Weighting Functions*

Ali al-Nowaihi, University of Leicester, UK
Sanjit Dhami, University of Leicester, UK

Working Paper No. 10/10
April 2010

Probability Weighting Functions*

Ali al-Nowaihi† **Sanjit Dhami**‡

16 April 2010

Abstract

In this paper we begin by stressing the empirical importance of non-linear weighting of probabilities, which expected utility theory (EU) is unable to accommodate. We then go on to outline three stylized facts on non-linear weighting that any alternative theory of risk must address. These are that people: overweight small probabilities and underweight large ones (S1); do not choose stochastically dominated options when such dominance is obvious (S2); ignore very small probabilities and code extremely large probabilities as one (S3). We then show that the concept of a probability weighting function (PWF) is crucial in addressing S1-S3. A PWF is not, however, a theory of risk. PWF's need to be embedded within some theory of risk in order to have significant predictive content. We ouline the two main alternative theories that are relevant in this regard: rank dependent utility (RDU) and cumulative prospect theory (CP). RDU and CP explain S1,S2 but not S3. We conclude by outlining the recent proposal for composite prospect theory (CPP) that uses the composite Prelec probability weighting function (CPF). CPF is axiomatically founded, and is flexible and parsimonious. CPP can explain all three stylized facts S1,S2,S3.

*Dr. Sanjit Dhami would like to acknowledge a one semester study leave provided by the University of Leicester. He also wishes to thank the International School of Economics (ISET) at Tbilisi State University, Georgia for their excellent hospitality and academic environment during the time that he was a visiting Professor in the spring of 2010.

†Department of Economics, University of Leicester, University Road, Leicester. LE1 7RH, UK. Phone: +44-116-2522898. Fax: +44-116-2522908. E-mail: aa10@le.ac.uk.

‡Department of Economics, University of Leicester, University Road, Leicester. LE1 7RH, UK. Phone: +44-116-2522086. Fax: +44-116-2522908. E-mail: Sanjit.Dhami@le.ac.uk.

Contents

Under expected utility theory (EU) decision makers weight probabilities linearly. However, the evidence suggests that decision makers weight probabilities in a non-linear manner. Consider, for instance, the following example from Kahneman and Tversky (1979, p. 283). Suppose that one is compelled to play Russian roulette. One would be willing to pay much more to reduce the number of bullets from one to zero than from four to three. However, in each case, the reduction in probability of a bullet firing is 1/6 and, so, under EU, the decision maker should be willing to pay the same amount. One possible explanation is that decision makers do not weight probabilities in a linear manner as under EU. There is also emerging evidence of the neuro-biological foundations for such behavior.[1] We now explore this idea more formally below.

1. Setting out the basics

1.1. Expected utility theory (EU)

Let $X = \{x_1, x_2, ..., x_n\}$ be a fixed, finite, set of real numbers, which can be interpreted as the possible monetary *outcomes* or possible wealth levels. We will assume throughout that the outcomes are arranged so that $x_1 < x_2 < ... < x_n$. The assumption that X is a finite set is not restrictive in practice because finite sets can be extremely large. The decision maker has a set of feasible actions, A. Any action in A induces a probability distribution $(p_1, p_2, ..., p_n)$, $p_i \geq 0$ and $\sum_{i=1}^{n} p_i = 1$, over the real numbers $x_1, x_2, ..., x_n$.

We then write a *lottery*, L, as

$$L = (x_1, p_1; x_2, p_2; ...; x_n, p_n), \tag{1.1}$$

with the interpretation that outcome x_i occurs with probability p_i. Let $\mathcal{L}$ be the set of such lotteries. How should the decision maker decide which action in the set A to choose? EU postulates the existence of an expected utility functional, $EU : \mathcal{L} \rightarrow \mathbb{R}$. Under well known assumptions, this functional takes the following form

$$EU(L) = \sum_{i=1}^{n} p_i u(x_i), \tag{1.2}$$

where $u(x_i)$ is the utility of the outcome x_i. EU predicts that the decision maker will choose that action (or lottery) which leads to the highest real number according to the functional, (1.2). The axioms required in the derivation of (1.2) can be found in many standard references, e.g., Fishburn (1982), MasCollel, Whinston and Green (1995), and Varian (1992). A key axiom that is relevant for us is the *independence axiom*.

Definition 1 *(Independence axiom): Suppose that* $\preceq$ *is a preference relation defined over the set of lotteries. Then for all lotteries* L_1, L_2, L, *and all* $p \in (0, 1]$, $L_1 \preceq L_2 \Leftrightarrow (L_1, p; L, 1-p) \preceq (L_2, p; L, 1-p)$.

[1] See, for instance, Berns et al (2008).

The independence axiom is often empirically violated. Alternatives to EU that we consider below, have in the main, relaxed the independence axiom. Two main features of EU stand out. (1) There is additive separability across outcomes, an assumption that EU shares with many alternative theories. (2) The objective function is linear in probabilities.

1.2. The Allais paradox and setting out the stylized facts

Most alternatives to EU relax the second feature, i.e., linearity in probabilities. The interested reader can consult Kahneman and Tversky (2000) and Starmer (2000) for a more detailed set of examples. We consider here the following example (problems 3 and 4) from Kahneman and Tversky (1979), based on the famous experiments of Allais.

Problem 3: 95 experimental subjects were asked to choose among the following two lotteries:

$$A = (4000, 0.8; 0, 0.2) \text{ or } B = (3000, 1)$$

20% of the subjects choose A while the remaining 80% choose B

Problem 4: 95 experimental subjects were asked to choose among the following two lotteries:

$$C = (4000, 0.2; 0, 0.8) \text{ or } D = (3000, 0.25; 0.0.75)$$

65% of the subjects choose C while the remaining 35% choose D.

Notice that Problem 4 differs from Problem 3 only in that the probabilities in the latter are scaled down by a quarter. Hence, given that probabilities are weighted linearly under EU, if A is preferred to B then C should be preferred to D. Conversely if B is preferred to A then D should be preferred to C. A simple substitution in (1.2) confirms this. But this pattern of preferences is violated by the experimental evidence. Most decision makers prefer B to A and C to D.

One possibility that can potentially explain the Allais paradox is that decision makers violate the linear weighting of probabilities in EU. To explore this reasoning further, suppose that the decision maker assigns decision weights $\pi(p) : [0,1] \to [0,1]$ that are non-linear, continuous functions of probabilities with $\pi(0) = 0$, $\pi(1) = 1$. The majority preference in Problem 3 can then be written as

$$\pi(0.8)u(4000) < \pi(1)u(3000) \Leftrightarrow \frac{\pi(0.8)}{\pi(1)} < \frac{u(3000)}{u(4000)}. \tag{1.3}$$

The majority preferences in Problem 4 can be written as:

$$\pi(0.2)u(4000) > \pi(0.25)u(3000) \Leftrightarrow \frac{\pi(0.2)}{\pi(0.25)} > \frac{u(3000)}{u(4000)} \tag{1.4}$$

From (1.3), (1.4) one can explain the preferences in Problems 3, 4 if the decision weights satisfy the following restriction:

$$\frac{\pi(0.8)}{\pi(1)} < \frac{\pi(0.2)}{\pi(0.25)}. \tag{1.5}$$

A property of decision weights, *subproportionality* (see, Kahneman and Tversky (1979), p. 282), implies that

$$\frac{\pi(pq)}{\pi(p)} \leq \frac{\pi(rpq)}{\pi(rp)}; r \leq 1. \tag{1.6}$$

Clearly subproportionality implies that (1.5) is satisfied (for $p = 1, q = 0.8, r = 0.25$). Subproportionality is, however, quite strong. As Kahneman and Tversky (1979) point out, it holds if and only if $\ln \pi$ is a convex function of $\ln p$.

Hence, decision weights provided an explanation for the sort of puzzles associated with, say, the Allais paradox. But this solution, had one drawback (in addition to the strong restrictions on $\pi(.)$). Decision makers who use such non-linear decision weights could choose stochastically dominated options even when such dominance is obvious, as the following example shows.

Example 1 *(Problems with non-linear weighting of probabilities) Consider the lottery* $(x, p; y, 1-p)$, $0 < p < 1$. *Let* $\pi(p)$ *denote decision weights such that* $\pi(p) : [0,1] \to [0,1]$. *Let* u *be a utility function over sure outcomes. Denote by* $\overline{U}$*, the non-expected utility functional, defined as* $\overline{U} = \pi(p) u(x) + \pi(1-p) u(y)$. *Let strict preferences induced by* $\overline{U}$ *over the set of all lotteries be denoted by* $\prec_{\overline{U}}$ *and the indifference relation be* $\sim_{\overline{U}}$.
(a) For the special case, $x = y$*, we get* $\overline{U} = [\pi(p) + \pi(1-p)] u(x)$. *In general, for non-EU theories,* $\pi(p) + \pi(1-p) \neq 1$. *Typically, for* $0 < p < 1$, $\pi(p) + \pi(1-p) < 1$, *hence* $(x, p; x, 1-p) \prec_{\overline{U}} (x; 1)$. *But any 'sensible' theory of risk should give* $(x, p; x, 1-p) \sim_{\overline{U}} (x; 1)$.
(b) Take $y = x + \epsilon$, $\epsilon > 0$. *Then, clearly,* $(x, 1) \prec_1 (x, p; x+\epsilon, 1-p)$*, where* $\prec_1$ *denotes first order stochastic dominance. Now,* $\overline{U}(x, p; x+\epsilon, 1-p) = \pi(p) u(x) + \pi(1-p) u(x+\epsilon)$. *Assuming* u *is continuous (and, as in (a),* $\pi(p) + \pi(1-p) < 1$*), we get* $\lim_{\epsilon \to 0} \overline{U}(x, p; x+\epsilon, 1-p) = \pi(p) u(x) + \pi(1-p) u(x) = [\pi(p) + \pi(1-p)] u(x) < U(x)$. *Hence, for sufficiently small* $\epsilon > 0$, $\overline{U}(x, p; x+\epsilon, 1-p) < U(x)$. *Thus,* $(x, 1) \prec_1 (x, p; x+\epsilon, 1-p)$ *but* $U(x, 1) \succ_{\overline{U}} U(x, p; x+\epsilon, 1-p)$. *Hence, monotonicity is violated.*

Based on the discussion so far, we summarize our first two stylized facts, S1, S2.

- S1 Decision makers weight probabilities in a non-linear manner. In particular, the evidence suggests that decision makers overweight low probabilities and underweight high probabilities. For further evidence, the reader could consult Kahneman and Tversky (2000) and Starmer (2000).

S2. Incorporation of non-linear probabilities into a prototype non-expected utility model (that induces, say, the preference ordering $\prec_{\overline{U}}$ in Example 1) is problematic. When decision weights transform objective probabilities, the decision maker might be lead to choose stochastically dominated options. In actual practice when such dominance is obvious, decision makers do not choose the dominated option. We leave open the question here of what decision makers do when stochastic dominance is not obvious. The interested reader can consult Dhami and al-Nowaihi (2010a) for further details.

We now highlight yet another important stylized fact on non-linear weighting of probabilities that has great significance at the end points of the probability interval $[0, 1]$.

S3. For events close to the boundary of the probability interval $[0, 1]$, extensive evidence, that we briefly review below, suggests the following. Decision makers (i) ignore events of extremely low probability and, (ii) treat extremely high probability events as certain.

Stylized facts S1 and S2 have been well documented in the literature. However, S3 is less well documented and theoretical models that incorporate it formally have only just about started to appear as discussion papers. For that reason, we discuss it in some detail here.

1.3. A brief note on prospect theory (PT)

Stylized fact S3 is given prominence in the Nobel prize winning work on *prospect theory* (PT), due to Kahneman and Tversky (1979). PT is a psychologically rich theory and it was the outcome of many years of experiments conducted by Kahneman, Tversky, and others. The psychological foundations of PT rest, in an important manner, on the distinction between an *editing* and an *evaluation/decision* phase. In PT, the term prospects can, for the time being, be thought to refer to our standard lotteries (we make the formal distinction between prospects and lotteries, later in our paper).

From our perspective, the most important and interesting aspect of the editing phase takes place when decision makers decide which improbable events to treat as impossible and which probable events to treat as certain. This is exemplified in the quote from Kahneman and Tversky (1979, p.282): "*the simplification of prospects can lead the individual to discard events of extremely low probability and to treat events of extremely high probability as if they were certain.*"

Suppose that we have a lottery $(x, p; y, 1-p)$ whose value to the decision maker is given by the *value function* $V = \pi(p)u(x) + \pi(1-p)u(y)$. In the editing phase, among other things, Kahneman and Tversky (1979) were interested in the decision weights, $\pi(p)$, assigned by individuals to very low and very high probability events. They drew $\pi(p)$ as in Figure 1.1. This decision function is discontinuous at both ends (arguably the most famous discontinuous function in Econometrica), reflecting the vexed issue of how decision makers behave over these ranges of probabilities.

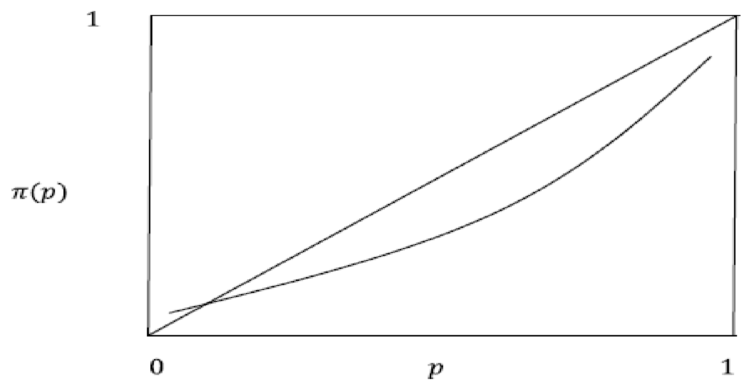


Figure 1.1: Ignorance at the endpoints. Source: Kahneman and Tversky (1979, p. 283)

Kahneman and Tversky's (1979) wrote the following (on p. 282-83) to expertly summarize the evidence on the end-points of the probability interval $[0, 1]$. "*The sharp drops or apparent discontinuities of* $\pi(p)$ *at the end-points are consistent with the notion that there is a limit to how small a decision weight can be attached to an event, if it is given any weight at all. A similar quantum of doubt could impose an upper limit on any decision weight that is less than unity. This quantal effect may reflect the categorical distinction between certainty and uncertainty. On the other hand, the simplification of prospects can lead the individual to discard events of extremely low probability and to treat events of extremely high probability as if they were certain. Because people are limited in their ability to comprehend and evaluate extreme probabilities, highly unlikely events are either ignored or overweighted, and the difference between high probability and certainty is either neglected or exaggerated. Consequently* $\pi(p)$ *is not well-behaved near the end-points.*"

After the prospects are 'psychologically cleaned' in the *editing phase*, the decision maker chooses the prospect with the highest numerical value assigned by the *value function,* V.

2. The importance of low probability events and problems for existing theory: A discussion

This section draws heavily on al-Nowaihi and Dhami (2010a,b) and Dhami and al-Nowaihi (2010b), which the interested reader could consult for the details. We present here some of the relevant evidence for stylized fact S3 by examining human behavior for low probability events from several economic and non-economic contexts. This is not an exhaustive list of such events but one that should suffice.

2.1. Insurance for low probability events

The insurance industry is of tremendous economic importance. Yet, despite impressive progress, existing theoretical models are unable to explain the stylized facts on the take-up of insurance for low probability events. The seminal study of Kunreuther et al. (1978) provides striking evidence of individuals buying inadequate, non-mandatory, insurance against low probability events (e.g., earthquake, flood and hurricane damage in areas prone to these hazards).

EU predicts that a risk-averse decision maker facing an actuarially fair premium will, in the absence of transactions costs, buy full insurance for all probabilities, however small. Kunreuther et al. (1978, chapter 7) presented subjects with varying potential losses with various probabilities, keeping the expected value of the loss constant. Subjects faced actuarially fair, unfair or subsidized premiums. In each case, they found that there is a point below which the take-up of insurance drops dramatically, as the probability of the loss decreases (and as the magnitude of the loss increases, keeping the expected loss constant). These results were shown to be robust to a very large number of perturbations and factors; see al-Nowaihi and Dhami (2010b) for the details.

Arrow's own reading of the evidence in Kunreuther et al. (1978) is that the problem is on the demand side rather than on the supply side. Arrow writes (Kunreuther et al., 1978, p.viii): "Clearly, a good part of the obstacle [to buying insurance] was the lack of interest on the part of purchasers." Kunreuther et al. (1978, p. 238) write: "Based on these results, we hypothesize that most homeowners in hazard-prone areas have not even considered how they would recover should they suffer flood or earthquake damage. Rather they treat such events as having a probability of occurrence sufficiently low to permit them to ignore the consequences." This behavior is in close conformity to the observations of Kahneman and Tversky (1979) outlined above.

2.2. Becker (1968) Paradox

A celebrated result, the Becker (1968) proposition, states that the most efficient way to deter a crime is to impose the '*severest possible penalty with the lowest possible probability*'. By reducing the probability of detection and conviction, society can economize on the costs of enforcement such as policing and trial costs. But by increasing the severity of the punishment (e.g., fines), which is not costly, the deterrence effect of the punishment is maintained. Indeed, under EU, and allowing for infinitely severe punishments, the Becker proposition implies that crime would be deterred completely, however small the probability of detection and conviction. Kolm (1973) memorably phrased this proposition as: *it is efficient to hang offenders with probability zero*. Empirical evidence, however, is not supportive of the Becker proposition; see, for example, Radelet and Ackers (1996), Levitt (2004), Polinsky and Shavell (2007: 422-23). For the details see Dhami and al-Nowaihi (2010b)

2.3. Evidence from jumping red traffic lights

Bar-Ilan (2000) and Bar-Ilan and Sacerdote (2001, 2004) provide near decisive evidence that the alternative explanations (except the implication that individuals ignore low probability events) cannot explain the Becker paradox; see Dhami and al-Nowaihi (2010b). They estimate that there are approximately 260,000 accidents per year in the USA caused by red-light running with implied costs of car repair alone of the order of $520 million per year. It stretches plausibility to assume that these are simply mistakes. In running red lights, there is a *small probability* of an accident, but, the consequences are self inflicted and potentially have infinite costs. Rephrased, running red traffic lights is similar to *hanging oneself with a very small probability*, which is similar to the Becker proposition.

Using Israeli data, Bar-Ilan (2000) calculated that the expected gain from jumping one red traffic is, at most, one minute (the length of a typical light cycle). Given the known probabilities they find that: "If a slight injury causes a loss greater or equal to 0.9 days, a risk neutral person will be deterred by that risk alone. However, the corresponding numbers for the additional risks of serious and fatal injuries are 13.9 days and 69.4 days respectively". To this must be added other costs arising from an accident. Dhami and al-Nowaihi (2010b) show that most of the attempts to explain the Becker paradox cannot work in this case because the punishment under jumping red traffic lights is self inflicted.

A far more natural explanation, along the lines of our framework, is that stylized fact S3 applies. Thus, red traffic light running is simply caused by some individuals ignoring (or seriously underweighting) the very low probability of an accident.

2.4. Driving and talking on car mobile phones

Consider the usage of mobile phones in moving vehicles. A user of mobile phones faces potentially infinite punishment (e.g., loss of one's and/or the family's life) with *low probability*, in the event of an accident. The Becker proposition applied to this situation would suggest that drivers of vehicles will not use mobile phones while driving or perhaps use hands-free phones, and so, avoid the self inflicted punishment. However, the evidence is to the contrary; see the Royal Society for the Prevention of Accidents (2005) and Pöystia et al. (2004). A natural explanation is the individuals simply ignore or substantially underweight the low probability of an accident as in stylized fact S3.

2.5. Other examples

People were reluctant to use seat belts prior to their mandatory use despite publicly available evidence that seat belts save lives. Prior to 1985, in the US, only 10-20% of motorists wore seat belts voluntarily, hence, denying themselves *self-insurance*; see Williams and Lund (1986). Car accidents may be perceived by individuals as *low probability events*, particularly if they are overconfident of their driving abilities. Overconfidence is supported from a wide range of contexts; see for instance Dhami and al-Nowaihi (2010a).

Even as evidence accumulated about the dangers of breast cancer (which has a *low unconditional probability*) women took up the offer of breast cancer examination, only sparingly. In the US, this changed after the greatly publicized events of the mastectomies of Betty Ford and Happy Rockefeller; see Kunreuther et al. (1978, p. xiii, p. 13-14).

2.6. Conclusion from these disparate contexts

Two main conclusions arise from the discussion in this section. First, human behavior for low probability events cannot be easily explained by the existing mainstream theoretical models of risk. al-Nowaihi and Dhami (2010a,b) show that EU and the associated auxiliary assumptions are unable to explain the stylized facts. And furthermore, the leading non-expected utility alternatives such as rank dependent utility (RDU), prospect theory (PT) and cumulative prospect theory (CP) make the problem even worse. Second, a natural explanation for these phenomena seems to be that individuals simply ignore or seriously underweight very low probability events (stylized fact S3).

There is some evidence of a bimodal perception of risks that could offer a potential explanation.[2] Some individuals focus more on the probability and others on the size of the loss. The former do not pay attention to losses that fall below a certain probability threshold, while for the latter, the size of the loss is relatively more salient. Hence, stylized fact S3 applies to the former set of individuals, which given the evidence, seem to predominate but are not the only types.

3. A look ahead and the notion of a probability weighting function (PWF)

Stylized facts S1, S2, S3 would seem to be the minimum requirements that a theory of risk should address. It would seem difficult to explain these stylized facts using linear weighting of probabilities, as in EU. Most alternatives to EU that use non-linear weighting of probabilities, such as rank dependent utility (RDU), prospect theory (PT) and cumulative prospect theory (CP) invoke the concept of a probability weighting function

[2] See Camerer and Kunreuther (1989) and for the evidence, see Schade et al (2001).

to incorporate stylized facts S1 and S2. None of these theories can incorporate all of the three stylized facts S1,S2,S3. We examine below emerging work due to al-Nowaihi and Dhami (2010a,b), and Dhami and al-Nowaihi (2010b) who propose *composite cumulative prospect theory* (CCP) and discuss applications that are able to address S1,S2,S3.

Remark 1 *It is important to note that a probability weighting function (PWF), by itself, is NOT a theory of risk. It needs to be embedded within other theories, such as RDU, PT, CP for it to have significant predictive content in concrete economic situations.*

We begin with some simple definitions.

Definition 2 *(Probability weighting function, PWF): By a probability weighting function we mean a strictly increasing function* $w(p) : [0, 1] \overset{onto}{\to} [0, 1]$.

A simple proof, that we omit, can be used to demonstrate the following properties of a PWF (for the proofs, see al-Nowaihi and Dhami, 2010a).

Proposition 1 *: A probability weighting function has the following properties:*
(a) $w(0) = 0$, $w(1) = 1$. *(b)* w *has a unique inverse,* w^{-1}, *and* w^{-1} *is also a strictly increasing function from* $[0, 1]$ *onto* $[0, 1]$. *(c)* w *and* w^{-1} *are continuous.*

The following definitions will prove useful to address stylized fact S3 and will also illustrate why standard PWF's are unable to address S3.

Definition 3 *: The function,* $w(p)$, *(a) infinitely-overweights infinitesimal probabilities, if* $\lim_{p\to 0} \frac{w(p)}{p} = \infty$, *and (b) infinitely-underweights near-one probabilities, if* $\lim_{p\to 1} \frac{1-w(p)}{1-p} = \infty$.

Definition 4 *: The function,* $w(p)$, *(a) zero-underweights infinitesimal probabilities, if* $\lim_{p\to 0} \frac{w(p)}{p} = 0$, *and (b) zero-overweights near-one probabilities, if* $\lim_{p\to 1} \frac{1-w(p)}{1-p} = 0$.

Definition 5 *: (a)* $w(p)$ *finitely-overweights infinitesimal probabilities, if* $\lim_{p\to 0} \frac{w(p)}{p} \in (1, \infty)$, *and (b)* $w(p)$ *finitely-underweights near-one probabilities, if* $\lim_{p\to 1} \frac{1-w(p)}{1-p} \in (1, \infty)$.

Definition 6 *: (a)* $w(p)$ *positively-underweights infinitesimal probabilities, if* $\lim_{p\to 0} \frac{w(p)}{p} \in (0, 1)$, *and (b)* $w(p)$ *positively-overweights near-one probabilities, if* $\lim_{p\to 1} \frac{1-w(p)}{1-p} \in (0, 1)$.

4. Addressing stylized fact S1

Data from experimental and field evidence typically suggest that decision makers exhibit an inverse S-shaped probability weighting over outcomes (stylized fact S1). Tversky and Kahneman (1992) propose the following probability weighting function, where the lower bound on γ comes from Rieger and Wang (2006).

Definition 7 : *The Tversky and Kahneman probability weighting function is given by*

$$w(p) = \frac{p^{\gamma}}{\left[p^{\gamma} + (1-p)^{\gamma}\right]^{\frac{1}{\gamma}}},\ 0.5 \le \gamma < 1,\ 0 \le p \le 1. \tag{4.1}$$

A simple proof leads to the following proposition.

Proposition 2 : *The Tversky and Kahneman (1992) probability weighting function (4.1) infinitely overweights infinitesimal probabilities and infinitely underweights near-one probabilities, i.e.,* $\lim_{p\to 0} \frac{w(p)}{p} = \infty$ *and* $\lim_{p\to 1} \frac{1-w(p)}{1-p} = \infty$*, respectively.*

Remark 2 *(Standard probability weighting functions): A large number of other probability weighting functions have been proposed, e.g., those by Gonzales and Wu (1999) and Lattimore, Baker and Witte (1992). Like the Tversky and Kahneman (1992) function, they all infinitely overweight infinitesimal probabilities and infinitely underweight near-one probabilities. We shall call these as the standard probability weighting functions. All these functions violate stylized fact S3.*

We now consider the most satisfactory PWF that helps to address stylized fact S1. This is the Prelec (1998) PWF, which is also a standard probability weighting function in the sense of remark 2. The Prelec (1998) PWF has the following merits: parsimony, consistency with much of the available empirical evidence (in the sense of stylized fact S1) and an axiomatic foundation.

Definition 8 *(Prelec, 1998): By the Prelec function we mean the probability weighting function* $w(p) : [0,1] \to [0,1]$ *given by*

$$w(0) = 0,\ w(1) = 1, \tag{4.2}$$

$$w(p) = e^{-\beta(-\ln p)^{\alpha}},\ 0 < p \le 1,\ \alpha > 0,\ \beta > 0. \tag{4.3}$$

The following Proposition is straightforward to check, so we omit the proof.

Proposition 3 : *The Prelec function (Definition 8) is a probability weighting function in the sense of Definition 2.*

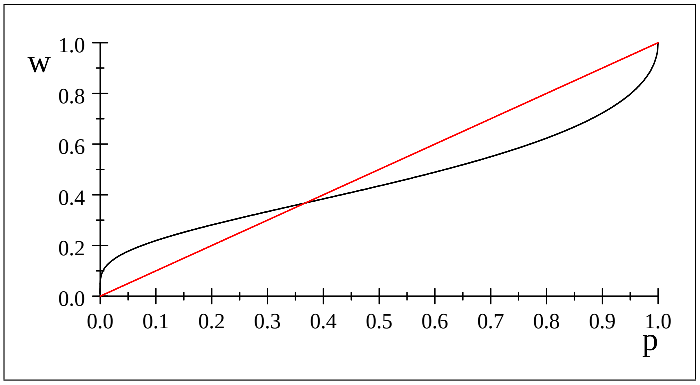


Figure 4.1: A plot of the Prelec (1988) function, $w(p) = e^{-(-\ln p)^{\frac{1}{2}}}$.

We plot the Prelec (1998) PWF in Figure 4.1 which is a plot of $w(p) = e^{-(-\ln p)^{0.5}}$, i.e., $\beta = 1$ and $\alpha = 0.5$ (see (4.3)) and the probability $p \in [0, 1]$.

Remark 3 *(Axiomatic foundations): Prelec (1998) gave an axiomatic derivation of (4.2) and (4.3) based on 'compound invariance', Luce (2001) provided a derivation based on 'reduction invariance' and al-Nowaihi and Dhami (2006) give a derivation based on 'power invariance'. Since the Prelec function satisfies all three, 'compound invariance', 'reduction invariance' and 'power invariance' are all equivalent. Note, in particular, that these derivations do not put any restrictions on α and β other than $\alpha > 0$ and $\beta > 0$.*

1. (Role of α) The parameter α controls the convexity/concavity of the Prelec function. If $\alpha < 1$, then the Prelec function is strictly concave for low probabilities but strictly convex for high probabilities, i.e., it is *inverse S-shaped*, as in $w(p) = e^{-(-\ln p)^{\frac{1}{2}}}$ ($\alpha = \frac{1}{2}$, $\beta = 1$), which is sketched in Figure 4.1, above. The converse holds if $\alpha > 1$. The Prelec function is then strictly convex for low probabilities but strictly concave for high probabilities, i.e., it is *S-shaped*. An examples is the curve $w(p) = e^{-(-\ln p)^2}$ ($\alpha = 2$, $\beta = 1$), sketched in Figure 4.2 as the light, lower, curve (the straight line in Figure 4.2 is the 45^o line).

2. (Role of β) Between the region of strict convexity ($w'' > 0$) and the region of strict concavity ($w'' < 0$), for each of the cases in Figures 4.1 and 4.2, there is a point of inflexion ($w'' = 0$). The parameter β in the Prelec function controls the location of the inflexion point relative to the 45^0 line. Thus, for $\beta = 1$, the point of inflexion is at $p = e^{-1}$ and lies on the 45^0 line, as in Figures 4.1 and 4.2 (light curve), above. However, if $\beta < 1$, then the point of inflexion lies above the 45^0 line, as in $w(p) = e^{-0.5(-\ln p)^2}$ ($\alpha = 2, \beta = 0.5$), sketched as the thicker, upper, curve in Figure 4.2. For this example, the fixed point, $w(p^*) = p^*$, is at $p^* \simeq 0.14$ but the point of inflexion, $w''(\widetilde{p}) = 0$, is at $\widetilde{p} \simeq 0.20$.

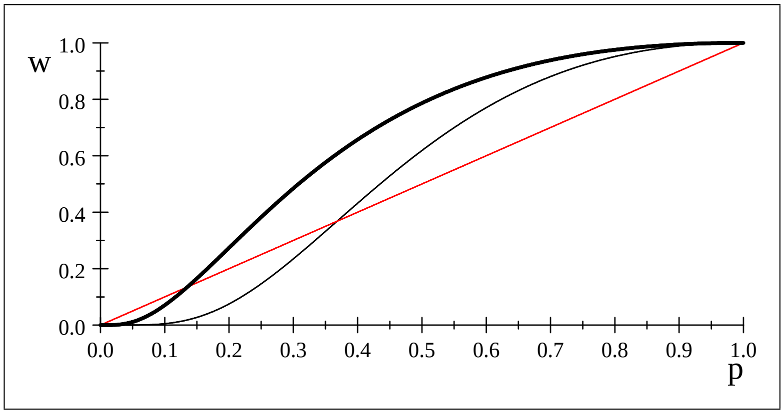


Figure 4.2: A plot of $w(p) = e^{-\frac{1}{2}(-\ln p)^2}$ and $w(p) = e^{-(-\ln p)^2}$.

Sometimes, the respective roles of α and β are also referred to as the *curvature* and *elevation* properties of a probability weighting function; see Gonzalez and Wu (1999) and Kilka and Weber (2001). Not all PWF's allow for a clear separation between curvature and elevation. This is particularly the case for PWF's that involve one parameter rather than two parameters. Some single parameter PWF's that have been proposed are: Karmarkar (1979), Röell (1987), Currim and Sarin (1989), Tversky and Kahneman (1992), Luce, Mellers and Chang (1993), Hey and Orme (1994), and Safra and Segal (1998). Among the two parameter PWF's are those by Goldstein and Einhorn (1987), Lattimore, Baker and Witte (1992) and Prelec (1998).

The full set of possibilities for the Prelec (1998) function is established by the following two propositions, which should be self-evident in the light of the discussion above. For the detailed proofs, the reader is referred to al-Nowaihi and Dhami (2010a).

Proposition 4 *: For $\alpha = 1$, the Prelec probability weighting function (Definition 8) takes the form $w(p) = p^{\beta}$, is strictly concave if $\beta < 1$ but strictly convex if $\beta > 1$. In particular, for $\beta = 1$, $w(p) = p$ (as under expected utility theory).*

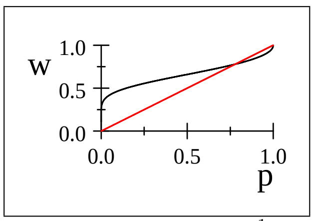
Proposition 5 *: Suppose $\alpha \neq 1$, then the Prelec PWF (Definition 8) has the following properties.*
(a) There are three fixed points, at respectively, 0, $p^* = e^{-\left(\frac{1}{\beta}\right)^{\frac{1}{\alpha-1}}}$ and 1.
(b) There is a unique inflexion point, $\widetilde{p} \in (0,1)$ at which $w''(\widetilde{p}) = 0$.
(c) If $\alpha < 1$, the Prelec function is strictly concave for $p < \widetilde{p}$ and strictly convex for $p > \widetilde{p}$ (inverse S-shaped). If $\alpha > 1$, then the converse holds
(d) For the cases $\beta < 1$, $\beta = 1$, $\beta > 1$, the respective inflexion points, $\widetilde{p}$, lie above ($\widetilde{p} < w(\widetilde{p})$), on ($\widetilde{p} = w(\widetilde{p})$) and below ($\widetilde{p} > w(\widetilde{p})$) the 45^0 line.

Table 1, below, exhibits the various cases established by Proposition 5.

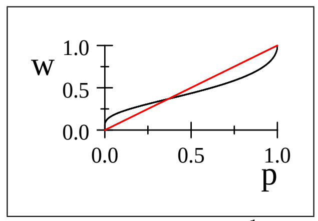
	$\beta < 1$	$\beta = 1$	$\beta > 1$
$\alpha < 1$	inverse S-shape $\widetilde{p} < w(\widetilde{p})$	inverse S-shape $\widetilde{p} = w(\widetilde{p})$	inverse S-shape $\widetilde{p} > w(\widetilde{p})$
$\alpha = 1$	strictly concave $p < w(p)$	$w(p) = p$	strictly convex $p > w(p)$
$\alpha > 1$	S-shape $\widetilde{p} < w(\widetilde{p})$	S-shape $w(\widetilde{p}) = \widetilde{p}$	S-shape $\widetilde{p} > w(\widetilde{p})$

Table 2, below, gives representative graphs of the Prelec function, $w(p) = e^{-\beta(-\ln p)^{\alpha}}$, for each of the cases in Table 1.

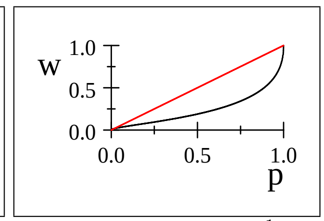
$\beta = \frac{1}{2}$ $\qquad$ $\beta = 1$ $\qquad$ $\beta = 2$



$w(p) = e^{-\frac{1}{2}(-\ln p)^{\frac{1}{2}}}$

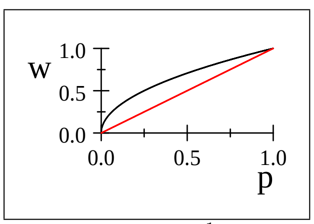


$w(p) = e^{-(-\ln p)^{\frac{1}{2}}}$.

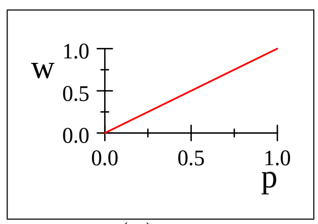


$w(p) = e^{-2(-\ln p)^{\frac{1}{2}}}$

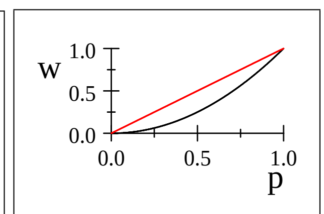
$\alpha = \frac{1}{2}$



$w(p) = p^{\frac{1}{2}}$

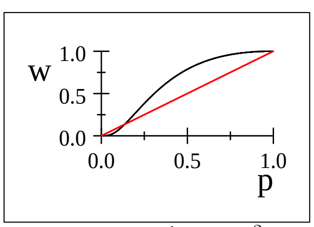


$w(p) = p$

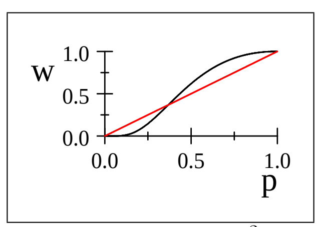


$w(p) = p^2$

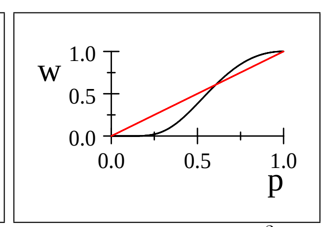
$\alpha = 1$



$w(p) = e^{-\frac{1}{2}(-\ln p)^2}$.



$w(p) = e^{-(-\ln p)^2}$.



$w(p) = e^{-2(-\ln p)^2}$

$\alpha = 2$

Table 2: Representative graphs of $w(p) = e^{-\beta(-\ln p)^{\alpha}}$.

Corollary 1 : *Suppose $\alpha \neq 1$. Then $\widetilde{p} = p^* = e^{-1}$ (i.e., the point of inflexion and the fixed point, coincide) if, and only if, $\beta = 1$. If $\beta = 1$, then:*
(a) If $\alpha < 1$, then w is strictly concave for $p < e^{-1}$ and strictly convex for $p > e^{-1}$ (inverse-S shape, see Figure 4.1).

(b) If $\alpha > 1$, then w is strictly convex for $p < e^{-1}$ and strictly concave for $p > e^{-1}$ (S shape, see Figure 4.2).

In Figure 4.1 (and first row in Table 2), where $\alpha < 1$, note that the slope of $w(p)$ becomes very steep near $p = 0$. By contrast, in figure 4.2 (and last row in Table 2), where $\alpha > 1$, the slope of $w(p)$ becomes very gentle near $p = 0$. This is established by the following proposition, which will be important for us to address stylized fact S3.

Proposition 6 : *(a) For $\alpha < 1$ the Prelec function (Definition 8): (i) infinitely-overweights infinitesimal probabilities, i.e., $\lim_{p\to 0} \frac{w(p)}{p} = \infty$, and (ii) infinitely underweights near-one probabilities, i.e., $\lim_{p\to 1} \frac{1-w(p)}{1-p} = \infty$ (Prelec, 1998, p505); see Definition 3 and Figure 4.1.*
(b) For $\alpha > 1$ the Prelec function: (i) zero-underweights infinitesimal probabilities, i.e., $\lim_{p\to 0} \frac{w(p)}{p} = 0$, and (ii) zero-overweights near-one probabilities, i.e., $\lim_{p\to 1} \frac{1-w(p)}{1-p} = 0$; see Definition 4 and figure 4.2.

According to Prelec (1998, p505), the infinite limits in Proposition 6a capture the qualitative change as we move from certainty to probability and from impossibility to improbability. On the other hand, they contradict stylized fact S3, i.e., the observed behavior that people ignore events of very low probability and treat very high probability events as certain; see, e.g., Kahneman and Tversky (1979). These specific problems are avoided for $\alpha > 1$. However, for $\alpha > 1$, the Prelec function is S-shaped, see Proposition 5(d) and Figure 4.2. This, however is in conflict with stylized fact S1.

4.1. Axiomatic derivations of Prelec's PWF

The probability weighting function with the strongest empirical support appears to be that of Prelec (1998). To quote from Stott (2006, p102): "the most predictive version of [cumulative prospect theory] has a power value curve, a single parameter risky weighting function due to Prelec (1998), and a Logit stochastic process." It was also the first axiomatically derived probability weighting function. In this subsection we overview three derivations of Prelec's probability weighting function: the original one by Prelec (1998), based on *compound invariance*, and the later ones by Luce (2001) and al-Nowaihi and Dhami (2006), based on, respectively, *reduction invariance* and *power invariance*.

For the purposes of this section, we shall assume that $0 \in X$, the set of outcomes, and we shall restrict ourselves to the special class of lotteries defined as follows.

Definition 9 : *Let $\mathbf{S} \subset \mathcal{L}$ be the subset of all lotteries of the forms (x), $(x,\ p_1)$, $((x,\ p_1),\ p_2)$ and $(((x,\ p_1),\ p_2),\ p_3)$, where $x \in X$ and $p_1, p_2, p_3 \in [0, 1]$.*

To simplify notation, we shall refer to $((x,\ p_1),\ p_2)$ and $(((x,\ p_1),\ p_2),\ p_3)$ by $(x;\ p_1,\ p_2)$ and $(x;\ p_1,\ p_2,\ p_3)$, respectively. Thus, (x) is the lottery whose outcome is x for sure, $(x;\ p_1)$ is the lottery whose outcomes are x with probability p_1 and 0 with probability $1-p_1$, $(x;\ p_1,\ p_2)$ is the lottery whose outcomes are $(x;\ p_1)$ with probability p_2 and 0 with probability $1-p_2$ and $(x;\ p_1,\ p_2,\ p_3)$ is the lottery whose outcomes are $(x;\ p_1,\ p_2)$ with probability p_3 and 0 with probability $1-p_3$.

Given a strictly increasing function, $u : X \to \mathbb{R}$, and a probability weighting function, w, we can extend u to a function, $U : \mathbf{S} \to \mathbb{R}$, by the following definition.

Definition 10 *:* $U(x) = u(x)$, $U(x;\ p_1) = w(p_1)U(x)$, $U(x;\ p_1,\ p_2) = w(p_2)U(x;\ p_1)$ *and* $U(x;\ p_1,\ p_2,\ p_3) = w(p_3)U(x;\ p_1,\ p_2)$.

Definition 11 *: Let* $\preceq$ *be the order on* $\mathbf{S}$ *induced by* U*, i.e., for all* $L_1, L_2 \in \mathbf{S}$, $L_1 \preceq L_2 \Leftrightarrow U(L_1) \leq U(L_2)$.

We are now in a position to introduce three apparently different axioms that, however, all lead to the Prelec probability weighting function.

Definition 12 *(Prelec, 1998): The preference relation,* $\preceq$*, satisfies* compound invariance *if, for all outcomes* $x,\ y,\ x',\ y' \in X$*, probabilities* $p,\ q,\ r,\ s \in [0,1]$ *and integers* $n \geq 1$*, the following holds. If* $(x,p) \sim (y,q)$ *and* $(x,r) \sim (y,s)$*, then* $(x',p^n) \sim (y',q^n)$ *implies* $(x',r^n) \sim (y',s^n)$.

Definition 13 *(Luce, 2001): The preference relation,* $\preceq$*, satisfies* reduction invariance *if, for all outcomes* $x \in X$*, probabilities* $p_1, p_2, q \in [0,1]$ *and* $\lambda \in \{2,3\}$*:* $(x;p_1,p_2) \sim (x;q) \Rightarrow \left(x;p_1^\lambda,p_2^\lambda\right) \sim \left(x;q^\lambda\right)$.

Definition 14 *(al-Nowaihi and Dhami, 2006): The preference relation,* $\preceq$*, satisfies* power invariance *if, for all outcomes* $x \in X$*, probabilities* $p,q \in [0,1]$ *and* $\lambda \in \{2,3\}$*:* $(x;\ p,p) \sim (x;q) \Rightarrow \left(x;\ p^\lambda,p^\lambda\right) \sim \left(x;q^\lambda\right)$ *and* $(x;\ p,p,p) \sim (x;q) \Rightarrow \left(x;\ p^\lambda,p^\lambda,p^\lambda\right) \sim \left(x;q^\lambda\right)$.

It is easy to check that for $w(p) = p$ *compound invariance*, *reduction invariance* and *power invariance* are all satisfied in EU. Hence, each of these constitutes a weakening (or generalization) of EU.

Proposition 7 *: A probability weighting function,* w*, is the Prelec probability weighting function if, and only if, the induced preference relation,* $\preceq$*, satisfies either* compound invariance*,* reduction invariance *or* power invariance*.*

Comparing Definitions 13 and 14, we see that *power invariance* is simpler in that it requires two probabilities (p, q) instead of three (p_1, p_2, q). On the other hand, *power invariance* requires two stages of compounding, instead of the single stage of *reduction invariance.* By contrast, *compound invariance* (Definition 12) involves no compounding but four probabilities and four outcomes. However, it follows from Proposition 7 that all three assumptions are equivalent. It is interesting to see that each of three a-priori different behavioral assumptions lead to Prelec's probability weighting function. We, thus, have a menu of testing options. Depending on the situation, one assumption may be more appropriate to test than the others. For further developments, the interested reader could further pursue Stott (2006) and Diecidue, Schmidt and Zank (2009).

5. Addressing stylized fact S2

There are two main ways of addressing S2. Either one uses *rank dependent expected utility theory* (RDU) or *cumulative prospect theory* (CP).

5.1. Rank dependent utility (RDU)

Quiggin (1982, 1993) for the first time, provided a coherent theory of behavior with non-linear weighting of probabilities. By 'coherent' is meant, among other things, that decision makers do not choose stochastically dominated options when such dominance is obvious. Quiggen's main insight was that it is not individual probabilities that should be transformed (which gave rise to the problem in Example 1) but rather, *cumulative probabilities.* When EU is applied to the transformed cumulative probabilities, we get what is now known as RDU.

RDU was a major advance in that it could easily solve some (though not all) of the paradoxes of EU. A major advantage of RDU over other behavioral decision theories is that the whole, and extensive, machinery of EU can be utilized (though applied to the transformed probabilities). By contrast, for other behavioral decisions theories, all these tools of economic analysis must be developed afresh. We now explain.

Let x be a random variable with distribution function $F(x)$. Then F is (1) non-decreasing, (2) right-continuous, and (3) $F(-\infty) = 0, F(\infty) = 1$. Let w be a PWF. Define Φ by

$$\Phi(x) = 1 - w(1 - F(x)). \tag{5.1}$$

Using the properties of a weighting function in Definition 2 it is simple to check that Φ also satisfies those properties. Notice that, Φ is a legitimate distribution function. Thus, we can define a new random variable, $\widetilde{x}$, by:

$$\Pr(\widetilde{x} \leq x) = \Phi(x) = 1 - w(1 - F(x)). \tag{5.2}$$

Consider the lottery $(x_1, p_1; x_2, p_2; ...; x_n, p_n)$, where $x_1 < x_2 < ... < x_n$; p_i is, of course, the probability that $x = x_i$. Let π_i be the probability that $\widetilde{x} = x_i$. Then

$$\begin{aligned} \pi_i &= \Phi(x_i) - \Phi(x_{i-1}) \\ &= [1 - w(1 - F(x_i))] - [1 - w(1 - F(x_{i-1}))] \\ &= w(1 - F(x_{i-1})) - w(1 - F(x_i)) \\ &= w\left(\Sigma_{j=i}^n p_j\right) - w\left(\Sigma_{j=i+1}^n p_j\right). \end{aligned} \tag{5.3}$$

The above considerations justify the following definitions (recall at this point, from the definition of a weighting function, that $w(0) = 0$ and $w(1) = 1$).

Definition 15 *: Consider the lottery $(x_1, p_1; x_2, p_2; ...; x_n, p_n)$, where $x_1 < x_2 < ... < x_n$. Let w be the probability weighting function. For RDU, the decision weights, π_i, are defined as follows.*
$\pi_n = w(p_n)$,
$\pi_{n-1} = w(p_{n-1} + p_n) - w(p_n)$,
...
$\pi_i = w\left(\Sigma_{j=i}^n p_j\right) - w\left(\Sigma_{j=i+1}^n p_j\right)$,
...
$\pi_1 = w\left(\Sigma_{j=1}^n p_j\right) - w\left(\Sigma_{j=2}^n p_j\right) = w(1) - w\left(\Sigma_{j=2}^n p_j\right) = 1 - w\left(\Sigma_{j=2}^n p_j\right)$.

Definition 16 *(rank dependent utility, RDU): Consider the lottery $(x_1, p_1; x_2, p_2; ...; x_n, p_n)$, where $x_1 < x_2 < ... < x_n$. Let w be the probability weighting function. Let π_i, $i = 1, 2, ..., n$, be given by Definition 15. The decision maker's rank dependent expected utility is given by*

$$U(x_1, p_1; x_2, p_2; ...; x_n, p_n) = \Sigma_{i=1}^n \pi_i u(x_i). \tag{5.4}$$

From Definition 15, we get that,

$$\pi_j \geq 0 \text{ and } \sum_{j=1}^n \pi_j = 1. \tag{5.5}$$

Remark 4 *: The result in (5.5) does not apply to the case of Example 1 because in the latter we transformed objective probabilities while in the former we transformed cumulative probabilities.*

Remark 5 *An alternative is to define $\Psi(x) = w(F(x))$. The decision weights would then be $\pi_i = w\left(\Sigma_{j=i}^n p_j\right) - w\left(\Sigma_{j=i-1}^n p_j\right)$. But the apparently more complex (5.1) to (5.3) are actually slightly more convenient.*

We now state the first main result of this section.

Proposition 8 : *A decision maker who uses RDU, as in definition 16, never chooses stochastically dominated options. Hence, such a decision maker uses non-linear transformations of cumulative probability but nevertheless is not subjected to the problem illustrated in Example 1. In other words, there is no problem with explaining stylized fact S2 for such a decision maker.*

The proof of Proposition 8 is not difficult but it is not too short either. For that reason we refer the interested reader to Quiggen (1982). A clearer proof can be found in Dhami and al-Nowaihi (2010a).

5.2. Cumulative prospect theory (CP)

The Nobel prize winning work of Kahneman and Tversky (1979), which is the second most cited paper in economics, suffered from the problem illustrated in Example 1. Namely, that decision makers could choose stochastically dominated options, even when such dominance was obvious. This did not go down too well with the profession. However, Quiggen's work on RDU provided Kahneman and Tversky with the means to address this problem by using cumulative transformations of probability. This, they duly did in their cumulative version of prospect theory, which they called as *cumulative prospect theory* (CP); see Tversky and Kahneman (1992); see also Starmer and Sugden (1989).

However, Tversky and Kahneman (1992) paid a heavy price for this desirable feature. Recalling the discussion in section 1.3 above, *CP dropped the editing phase altogether*, hence, also, giving up the psychological richness of PT. Tversky and Kahneman had no choice but to do so in order to address S2. In PT, the decision weights (see Figure 1.1) are discontinuous at both ends. However, in order to incorporate S2 using Quiggen's insights on cumulative transformations of probability they needed a continuous function. This, in turn, meant that editing of prospects, which created discontinuities at both ends had to be dispensed with.

The most persuasive contributory role played by a PWF, in terms of addressing a wide range of problems in economics arises when it is embedded within CP. We now outline, briefly, the basics of CP. Consider a lottery of the form

$$L = \left(y_{-m}, p_{-m}; y_{-m+1}, p_{-m+1}; ...; y_{-1}, p_{-1}; y_0, p_0; y_1, p_1; y_2, p_2; ...; y_n, p_n\right),$$

where $y_{-m} < ... < y_{-1} < y_0 < y_1 < ... < y_n$ are the outcomes or the *final* positions of aggregate wealth and $p_{-m}, ..., p_{-1}, p_0, p_1, ...p_n$ are the corresponding objective probabilities, such that $\sum_{i=-m}^{n} p_i = 1$ and $p_i \geq 0$. In CP, decision makers derive utility from wealth relative to a reference point for wealth, y_0, which could be initial wealth, status-quo wealth, average wealth, desired wealth, rational expectations of future wealth etc. depending on the context.

Definition 17 *(Lotteries in incremental form) Let* $x_i = y_i - y_0, i = -m, -m+1, ..., n$ *be the increment in wealth relative to* y_0 *and* $x_{-m} < ... < x_0 = 0 < ... < x_n$*. Let the restriction on probabilities be* $\Sigma_{i=-m}^{n} p_i = 1$*,* $p_i \geq 0$*,* $i = -m, -m+1, ..., n$*. Then, a lottery is presented in incremental form if it is represented as:*

$$L = (x_{-m}, p_{-m}; ...; x_{-1}, p_{-1}; x_0, p_0; x_1, p_1; ...; x_n, p_n)\,. \tag{5.6}$$

Definition 18 *(Lotteries and prospects): Lotteries that are represented in incremental form are known as prospects.*

Definition 19 *(Set of Lotteries): Denote by* $\mathcal{L}_P$ *the set of all prospects of the form given in (5.6) subject to the restrictions in definition 17.*

Definition 20 *(Domains of losses and gains): The decision maker is said to be in the domain of gains if* $x_i > 0$ *and in the domain of losses if* $x_i < 0$*.* x_0 *lies neither in the domain of gains nor in the domain of losses.*

5.2.1. The utility function in CP

Definition 21 *(Tversky and Kahneman, 1979). A utility function,* $v(x)$*, is a continuous, strictly increasing, mapping* $v : \mathbb{R} \to \mathbb{R}$ *that satisfies:*
1. $v(0) = 0$ *(reference dependence).*
2. $v(x)$ *is concave for* $x \geq 0$ *(declining sensitivity for gains).*
3. $v(x)$ *is convex for* $x \leq 0$ *(declining sensitivity for losses).*
4. $-v(-x) > v(x)$ *for* $x > 0$ *(loss aversion).*

Tversky and Kahneman (1992) propose the following utility function:

$$v(x) = \begin{cases} x^{\gamma} & \text{if} \quad x \geq 0 \\ -\lambda(-x)^{\theta} & \text{if} \quad x < 0 \end{cases} \tag{5.7}$$

where γ, θ, λ are constants. The coefficients of the power function satisfy $0 < \gamma < 1, 0 < \theta < 1$. $\lambda > 1$ is known as the *coefficient of loss aversion*. Tversky and Kahneman (1992) assert (but do not prove) that the axiom of *preference homogeneity* ($(x, p) \sim y \Rightarrow (kx, p) \sim ky$) generates this value function. al-Nowaihi et al. (2008) give a formal proof, as well as some other results (e.g. that γ is necessarily identical to θ). Tversky and Kahneman (1992) estimated that $\gamma \simeq \theta \simeq 0.88$ and $\lambda \simeq 2.25$. The reader can visually check the properties listed in definition 21 for the utility function, (5.7), plotted in figure 5.1 for the case:

$$\left\{ v(x) = \begin{array}{cc} \sqrt{x} & \text{if} \quad x \geq 0 \\ -2.5\sqrt{-x} & \text{if} \quad x < 0 \end{array} \right. \tag{5.8}$$

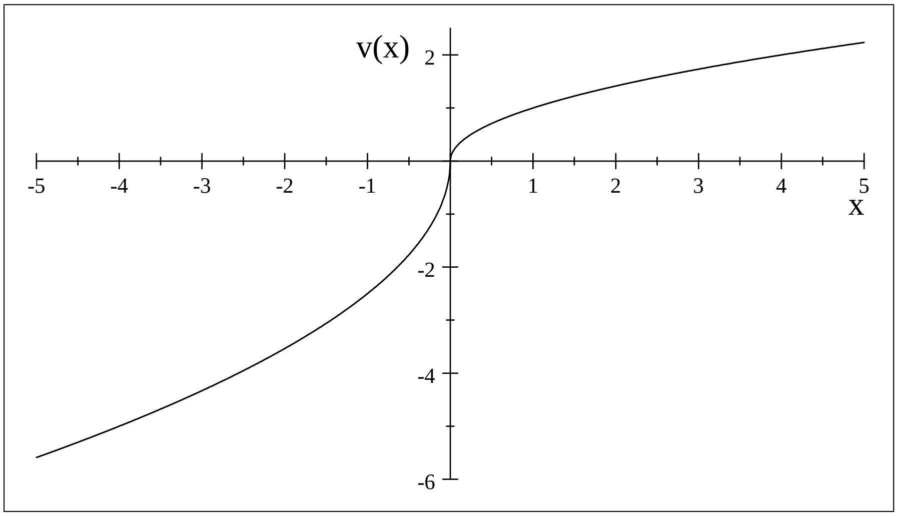


Figure 5.1: The utility function under CCP for the case in (5.8)

5.2.2. Construction of decision weights under CP

Let $w(p)$ be a PWF, such as the Prelec PWF in Definition 8. We could have different weighting functions for the domain of gains and losses, respectively, $w^{+}(p)$ and $w^{-}(p)$. However, we make the empirically founded assumption that $w^{+}(p) = w^{-}(p)$; see Prelec (1998).

Definition 22 *(Tversky and Kahneman, 1992). For CP, the decision weights, π_i, are defined as follows:*

Domain of Gains	*Domain of Losses*
$\pi_n = w(p_n)$	$\pi_{-m} = w(p_{-m})$
$\pi_{n-1} = w(p_{n-1} + p_n) - w(p_n) \ldots$	$\pi_{-m+1} = w(p_{-m} + p_{-m+1}) - w(p_{-m}) \ldots$
$\pi_i = w\left(\Sigma_{j=i}^{n} p_j\right) - w\left(\Sigma_{j=i+1}^{n} p_j\right) \ldots$	$\pi_j = w\left(\Sigma_{i=-m}^{j} p_i\right) - w\left(\Sigma_{i=-m}^{j-1} p_i\right) \ldots$
$\pi_1 = w\left(\Sigma_{j=1}^{n} p_j\right) - w\left(\Sigma_{j=2}^{n} p_j\right)$	$\pi_{-1} = w\left(\Sigma_{i=-m}^{-1} p_i\right) - w\left(\Sigma_{i=-m}^{-2} p_i\right)$

5.2.3. The objective function under prospect theory

As in EU, a decision maker using CP maximizes a well defined objective function, called the value function, which we now define.

Definition 23 *(The value function under CP) The value of the prospect, L_P, to the decision maker is given by*

$$V(L_P) = \Sigma_{i=-m}^{n} \pi_i v(x_i). \tag{5.9}$$

Note that the decision weights across the domain of gains and losses *do not* necessarily add up to 1. This contrasts with the case of RDU, in which there is no conception of different domains of gains and losses and the decision weights add up to one. To see this, from definition 22, we get that

$$\sum_{j=-m}^{n} \pi_j = w\left(\Sigma_{j=1}^{n} p_j\right) + w\left(\Sigma_{i=-m}^{-1} p_i\right) \neq 1 \tag{5.10}$$

If all outcomes were in the domain of gains then we get $\sum_{j=1}^{n} \pi_j = w\left(\Sigma_{j=1}^{n} p_j\right) = 1$ because $\Sigma_{j=1}^{n} p_j = 1$ and $w(1) = 1$ (as in RDU). If all outcomes were in the domain of losses then similarly $\sum_{j=-m}^{-1} \pi_j = w\left(\Sigma_{j=-m}^{-1} p_j\right) = 1$ because $\Sigma_{j=-m}^{-1} p_j = 1$ and $w(1) = 1$ (as in RDU). For the general case when there are some outcomes in the domain of gains and others in loss then, since $v(0) = 0$, the decision weight on the reference outcome, π_0, can be chosen arbitrarily. We have found it technically convenient to define $\pi_0 = 1 - \Sigma_{i=-m}^{-1}\pi_i - \Sigma_{i=1}^{n}\pi_i$, so that $\Sigma_{i=-m}^{n}\pi_i = 1$.

We now give the analogue of Proposition 8 for the case of CP and show that CP, too, can address S2. For the proof, see Dhami and al-Nowaihi (2010a).

Proposition 9 *: A decision maker who uses CP does not chooses stochastically dominated options. Hence, CP is able to address stylized fact S2.*

Remark 6 *: Using any of the standard probability weighting functions, CP (and RDU) can explain S1 but not S2.*

The PWF plays an important role in determining a rich set of attitudes towards risk under CP. Under EU, attitudes to risk are determined purely by the curvature of the utility function. So, given Definition 21(2),(3), it might be tempting to conclude that under CP the decision maker is risk averse in the domain of gains and risk loving in the domain of losses. This turns out not to be true because of the role played by the interaction of the PWF and the curvature of the utility function under CP, in determining attitudes to risk. The following four-fold pattern of risk preferences can be show under CP; see Kahneman and Tversky (2000). The decision maker is risk loving for small probabilities in the domain of gains and non-small probabilities in the domain of losses. He/she is also risk averse for non-small probabilities in the domain of gains and small probabilities in the domain of losses. The following example shows experimental evidence supporting these results.

Let C be the certainty equivalent of a prospect and E the expected value. As is well known from elementary microeconomics, for a risk averse decision maker, $C < E$ while for a risk loving decision maker $C > E$.

Example 2 *(The four fold pattern of attitudes to risk; Tversky and Kahneman, 1992).*

Probability	*Gain*	*Loss*
Low	$(100, 0.05; 0, 0.95); C = 14, E = 5$	$(-100, 0.05; 0, 0.95); C = -8, E = 5$
High	$(100, 0.95; 0, 0.05); C = 78, E = 95$	$(-100, 0.95; 0, 0.05); C = -84, E = 95$

In the table, $(x, p; 0, 1-p)$ denotes the lottery 'win x with probability p or 0 with probability $1-p$'. Comparing the values of C and E, we see that for low probabilities, there is risk seeking in the domain of gains but risk aversion in the domain of losses. For moderate or high probabilities, we see risk aversion in the domain of gains but risk seeking in the domain of losses. This pattern is hard to explain with EU or RDEU, but is easily explained by CP. Thus, the attitudes to risk in CP are more complex than under EU (or RDEU), and are the result of interaction between the shape of the value function, and the probability weighting function.

6. Addressing stylized fact S3

RDU and CP in conjunction with any of the standard PWF, say, the Prelec (1998) PWF are able to explain stylized facts S1 and S2. However, importantly, they are unable to address stylized fact S3. This has been an open problem since, at least, the paper by Kahneman and Tversky (1979).

al-Nowaihi and Dhami (2010a) make the ambitious proposal of combining the psychological richness of PT with the more satisfactory cumulative transformation of probabilities in CP. In other words, they combine PT and CP into a single theory, that they call *composite cumulative prospect theory* (CCP). If it aids intuition, CCP can be described as combining the editing and decision phases of PT (see section 1.3, above) into a single phase, while retaining cumulative transformations of probability, as in CP. CCP accounts for all three stylized facts S1, S2 and S3. It can explain everything that RDU and CP can, and addition more, that RDU and CP cannot (in particular, stylized fact S3).

From remark 6, the Prelec weighting function explains S1 but fails on S3. In order to implement CCP, al-Nowaihi and Dhami (2010a) introduce a modification to the Prelec (1998) PWF (see Figure 4.1) in a manner that is consistent with the empirical evidence. In other words, they eliminate the discontinuities at the end-points in Figure 1.1 with empirically-founded as well as axiomatically-founded behavior. They call their suggested modification as *composite Prelec weighting function* (CPF). Figure 6.1 sketches the CPF, which, in conjunction with either RDU or CP can potentially address S1, S2, S3.

In Figure 6.1, decision makers heavily underweight very low probabilities in the range $[0, p_1]$. Compare this to remark 2 for standard PWF's of which the Prelec (1998) function is an example. Akin to Kahneman and Tversky's (1979) editing phase, decision makers who use the weighting function in Figure 6.1 would typically ignore very low probability events

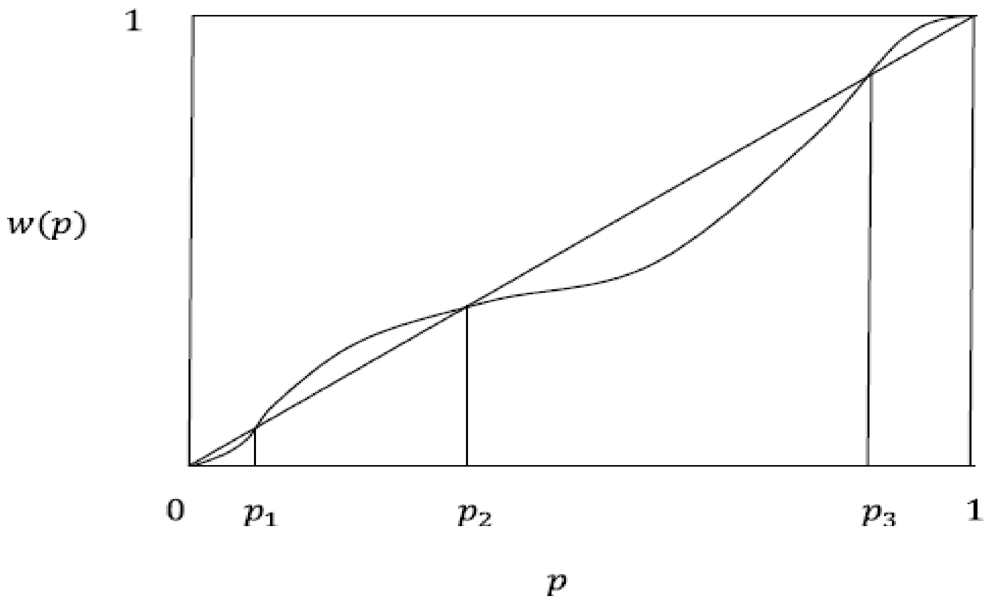


Figure 6.1: The composite Prelec weighting function (CPF).

by assigning low subjective weights to them. Hence, in conformity with the evidence (see Section 2) they are unlikely to be dissuaded from low-probability high-punishment crimes, reluctant to buy insurance for very low probability events (unless mandatory), reluctant to wear seat belts (unless mandatory), reluctant to participate in voluntary breast screening programs (unless mandatory) and so on. Similar comments apply to the probability range $[p_3, 1]$ except that events with these probabilities are overweighted as suggested by the evidence; see Kahneman and Tversky's (1979, p.282-83) quote, above. In the middle segment, $p \in [p_1, p_3]$, the probability weighting function in Figure 6.1 is identical to the Prelec function, and so addresses stylized fact S1.

For examples of the CPF, fitted to actual data, see al-Nowaihi and Dhami (2010a,b). Due to space limitations we restrict ourselves to a brief, formal, description of the CPF. This implements the general shape of the CPF in Figure 6.1. Define,

$$\underline{p} = e^{-\left(\frac{\beta}{\beta_0}\right)^{\frac{1}{\alpha_0 - \alpha}}};\ \overline{p} = e^{-\left(\frac{\beta}{\beta_1}\right)^{\frac{1}{\alpha_1 - \alpha}}}. \tag{6.1}$$

Essentially, the CPF in Figure 6.1 is described by segments from three different Prelec probability weighting functions. The first is defined over the range $0 < p \leq \underline{p}$, the second over the range $\underline{p} < p \leq \overline{p}$, and the third is defined over the range $\overline{p} < p \leq 1$.

Definition 24 *(al-Nowaihi and Dhami, 2010a): By the composite Prelec weighting func-*

tion we mean the probability weighting function $w : [0,1] \to [0,1]$ *given by*

$$w(p) = \begin{cases} 0 & if \quad p = 0 \\ e^{-\beta_0(-\ln p)^{\alpha_0}} & if \quad 0 < p \leq \underline{p} \\ e^{-\beta(-\ln p)^{\alpha}} & if \quad \underline{p} < p \leq \overline{p} \\ e^{-\beta_1(-\ln p)^{\alpha_1}} & if \quad \overline{p} < p \leq 1 \end{cases} \tag{6.2}$$

where $\underline{p}$ *and* $\overline{p}$ *are given by (6.1) and*

$$0 < \alpha < 1, \beta > 0; \alpha_0 > 1, \beta_0 > 0; \alpha_1 > 1, \beta_1 > 0, \beta_0 < 1/\beta^{\frac{\alpha_0 - 1}{1-\alpha}}, \beta_1 > 1/\beta^{\frac{\alpha_1 - 1}{1-\alpha}}. \tag{6.3}$$

Since a PWF (Definition 2) is one to one and onto over the interval, $[0,1]$, Definition 24 implies that $w(1) = 1$.

Proposition 10 *(al-Nowaihi and Dhami, 2010a): The composite Prelec function is a probability weighting function.*

The restrictions in (6.3) are required by the axiomatic derivations of the Prelec function (see section 4.1) and to ensure continuity of the CPF; see al-Nowaihi and Dhami (2010a) for the details.

Define p_1, p_2, p_3 that correspond to the notation used for the general shape of a CPF in Figure 6.1.

$$p_1 = e^{-\left(\frac{1}{\beta_0}\right)^{\frac{1}{\alpha_0 - 1}}}, \ p_2 = e^{-\left(\frac{1}{\beta}\right)^{\frac{1}{\alpha - 1}}}, \ p_3 = e^{-\left(\frac{1}{\beta_1}\right)^{\frac{1}{\alpha_1 - 1}}} \tag{6.4}$$

Proposition 11 *(al-Nowaihi and Dhami, 2010a): (a)* $p_1 < \underline{p} < p_2 < \overline{p} < p_3$*. (b)* $p \in (0, p_1) \Rightarrow w(p) < p$*. (c)* $p \in (p_1, p_2) \Rightarrow w(p) > p$*. (d)* $p \in (p_2, p_3) \Rightarrow w(p) < p$*. (e)* $p \in (p_3, 1) \Rightarrow w(p) > p$*.*

By Proposition 10, the CPF in (6.2), (6.3) is a PWF in the sense of Definition 2. By Proposition 11, a CPF overweights low probabilities, i.e., those in the range (p_1, p_2), and underweights high probabilities, i.e., those in the range (p_2, p_3). Thus, it accounts for stylized fact S1. But, in addition, and unlike all the standard probability weighting functions, it underweights probabilities near zero, i.e., those in the range $(0, p_1)$, and overweights probabilities close to one, i.e., those in the range $(p_3, 1)$ as required in S3.

Proposition 12 *(al-Nowaihi and Dhami, 2010a): The CPF (6.2):*
(a) zero-underweights infinitesimal probabilities, i.e., $\lim_{p \to 0} \frac{w(p)}{p} = 0$ *(Definition 4a),*
(b) zero-overweights near-one probabilities, i.e., $\lim_{p \to 1} \frac{1 - w(p)}{1-p} = 0$ *(Definition 4b).*

al-Nowaihi and Dhami (2010a) show that their proposed probability weighting function in Figure 6.1 is *axiomatic*, *parsimonious* and *flexible*. The axiomatic derivation uses the axiom of *local power invariance*, which is a variant of the axiom that al-Nowaihi and Dhami (2006) use in the proof of the Prelec PWF.

Definition 25 *(al-Nowaihi and Dhami, 2010a): Otherwise standard CP, when combined with a CPF is called composite cumulative prospect theory (CCP). Analogously, otherwise standard RDU, when combined with a CPF, is referred to as composite rank dependent utility (CRDU).*

al-Nowaihi and Dhami (2010a) prove the following proposition, whose intuition would by now be largely clear to the reader from our discussion of the CPF. Because probabilities in the middle ranges are weighted as in Prelec (1998), stylized fact S1 is explained. Because cumulative transformations of probability are undertaken in CCP and CRDU, S2 is explained. And because of the property of the CPF in Proposition 12, S3 is explained.

Proposition 13 *(al-Nowaihi and Dhami, 2010a): CCP and CRDU can explain S1, S2 and S3.*

Blavatskyy (2005) shows that the St. Petersberg paradox re-emerges under CP. al-Nowaihi and Dhami (2010a) show that the St. Petersberg paradox can be resolved under CCP mainly through the role played by the CPF. Rieger and Wang (2006) also propose a PWF that resolves the St. Petersberg paradox but it cannot explain stylized fact S3.

In comparison to CRDU, CCP, in addition, incorporates reference dependence, loss aversion and richer attitudes towards risk. Hence, it can explain everything that CRDU can, but the converse is false. Furthermore, because CCP explains S1, S2 and S3, while CP (and RDU) can only explain S1, S2, CCP can explain everything that CP (and RDU) can, but the converse is false. In light of these observations it is interesting to note the observation in Machina (2008) that "RDU is currently the most popular decision theory under risk."

References

[1] al-Nowaihi, A., Dhami, S., (2010a) Composite prospect theory: A proposal to combine prospect theory and cumulative prospect theory. University of Leicester Discussion Paper.

[2] al-Nowaihi, A., Dhami, S., (2010b) Insurance behavior for low probability events. University of Leicester Discussion Paper.

[3] al-Nowaihi, A., Dhami, S., (2006) A simple derivation of Prelec's probability weighting function. *The Journal of Mathematical Psychology* 50 (6), 521-524.

[4] al-Nowaihi, A., Dhami, S. and Bradley, I., (2008) 'The Utility Function Under Prospect Theory', *Economics Letters* 99, p.337–339.

[5] Bar Ilan, Avner (2000) “The Response to Large and Small Penalties in a Natural Experiment”, Department of Economics, University of Haifa, 31905 Haifa, Israel.

[6] Bar Ilan, Avner and Bruce Sacerdote (2001) “The response to fines and probability of detection in a series of experiments” National Bureau of Economic Research Working Paper 8638.

[7] Bar Ilan, Avner and Bruce Sacerdote (2004) “The response of criminals and noncriminals to fines” *Journal of Law and Economics*, Vol. XLVII, 1-17.

[8] Becker, Gary (1968) “Crime and Punishment: an Economic Approach, ” *Journal of Political Economy*, 76, 169-217.

[9] Berns, G. S., Capra, C. M., Chappelow, J., Moore, and Noussair, C. (2008) Nonlinear neurobiological probability weighting functions for aversive outcomes. *Neuroimage*, 39: 2047-57.

[10] Blavatskyy, P.R., (2005) Back to the St. Petersburg paradox? *Management Science*, 51, 677-678.

[11] Camerer, Colin F. (2000) Prospect theory in the wild: Evidence from the field. In D. Kahneman and A. Tversky (eds.), Choices, Values, and Frames, Cambridge University Press; Cambridge.

[12] Camerer, C. and Kunreuther, H. (1989) Experimental markets for insurance. Journal of Risk and Uncertainty, 2, 265-300.

[13] Currim, Imran S. and Rakesh K. Sarin (1989), “Prospect Versus Utility,” *Management Science* 35, 22-41

[14] Diecidue Enrico, Ulrich Schmidt and Horst Zank (2009) "Parametric Weighting Functions," *Journal of Economic Theory* 144, 1102-1118.

[15] Dhami, S. and al-Nowaihi, A. (2010a) Behavioral Economics: Theory, Evidence and Applications, Book manuscript in progress.

[16] Dhami, S. and al-Nowaihi, A. (2010b) The Becker paradox reconsidered through the lens of behavioral economics. University of Leicester Discussion Paper.

[17] Fishburn, Peter C. (1982) The foundations of expected utility. D. Reidel Publishing Company. Dordrecht: Holland/Boston: U.S.A. London: England.

[18] Goldstein, W. M., and Einhorn, H. J. (1987), “Expression Theory and the Preference Reversal Phenomena,” *Psychological Review* 94, 236-254.

[19] Gonzalez, R., Wu, G., (1999) On the shape of the probability weighting function. *Cognitive Psychology* 38, 129-166.

[20] Hey, John D. and Chris Orme (1994), "Investigating Generalizations of Expected Utility Theory Using Experimental Data," *Econometrica* 62, 1291-1326.

[21] Kahneman, D., and Tversky, A. (Eds.). (2000). Choices, values and frames. New York: Cambridge University Press.

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

[23] Karmarkar, Uday S. (1979), "Subjectively Weighted Utility and the Allais Paradox," *Organizational Behavior and Human Performance* 24, 67-72.

[24] Kilka, Michael and Martin Weber (2001), "What Determines the Shape of the Probability Weighting Function under Uncertainty?" *Management Science* 47, 1712-1726.

[25] Kolm, Serge-Christophe (1973) "A Note on Optimum Tax Evasion," *Journal of Public Economics*, 2, 265-270.

[26] Kunreuther, H., Ginsberg, R., Miller, L., Sagi, P., Slovic, P., Borkan, B., Katz, N., (1978) Disaster insurance protection: Public policy lessons, Wiley, New York.

[27] Lattimore, J. R., Baker, J. K., Witte, A. D. (1992) The influence of probability on risky choice: A parametric investigation. *Journal of Economic Behavior and Organization* 17, 377-400.

[28] Levitt, S. D. (2004) "Understanding why crime fell in the 1990's: Four factors that explain the decline and six that do not." *Journal of Economic Perspectives*, 18: 163-190.

[29] Luce, R. D. (2001) Reduction invariance and Prelec's weighting functions. *Journal of Mathematical Psychology* 45, 167-179.

[30] Luce, R. Duncan, Barbara A. Mellers, and Shi-Jie Chang (1993), "Is Choice the Correct Primitive? On Using Certainty Equivalents and Reference Levels to Predict Choices among Gambles," *Journal of Risk and Uncertainty* 6, 115-143.

[31] Machina, M. J. (2008) Non-expected utility theory. in Durlaf, S. N. and Blume, L. E. (eds.) The New Palgrave Dictionary of Economics, 2nd Edition, Macmillan (Basingstoke and New York).

[32] Mas-Collel, Andreu, Michael D. Whinston, and Jerry R. Green. (1995) Microeconomic Theory, New. York: Oxford Univ. Press.

[33] Polinsky, Mitchell and Steven Shavell (2007) "The theory of public enforcement of law." Chapter 6 in Polinsky, Mitchell and Steven Shavell (eds.) Handbook of Law and Economics, Vol. 1, Elsevier.

[34] Pöystia, L., Rajalina, S., and Summala, H., (2005) "Factors influencing the use of cellular (mobile) phone during driving and hazards while using it." *Accident Analysis & Prevention*, 37: 47-51.

[35] Prelec, D., (1998) The probability weighting function. Econometrica 60, 497-528.

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

[37] Quiggin, J. (1993) Generalized Expected Utility Theory, Kluwer Academic Publishers.

[38] Radelet, M. L. and R. L. Ackers, (1996) "Deterrence and the Death Penalty: The Views of the Experts." Journal of Criminal Law and Criminology 87 : 1-16.

[39] Rieger, M.O., Wang, M. (2006) Cumulative prospect theory and the St. Petersburg paradox. Economic Theory 28, 665-679.

[40] Röell, Ailsa (1987), "Risk Aversion in Quiggin and Yaari's Rank-Order Model of Choice under Uncertainty," *Economic Journal* 97, 143-160.

[41] Royal Society for the Prevention of Accidents (2005) 'The risk of using a mobile phone while driving'. Full text of the report can be found at www.rospa.com.

[42] Safra, Zvi and Uzi Segal (1998), "Constant Risk Aversion," *Journal of Economic Theory* 83, 19-42.

[43] Schade, C., Kunreuther, H., and Kaas, K. P. (2001) Low-Probability Insurance: Are Decisions Consistent with Normative Predictions? mimeo. University of Humboldt.

[44] Starmer, C. (2000) "Developments in Non-expected Utility Theory: The Hunt for a Descriptive Theory of Choice under Risk," *Journal of Economic Literature*, 38: 332-382.

[45] Starmer, Chris and Robert Sugden (1989), "Probability and Juxtaposition Effects: An Experimental Investigation of the Common Ratio Effect," J*ournal of Risk and Uncertainty* 2, 159-178.

[46] Stott, H.P. (2006) Choosing from cumulative prospect theory's functional menargerie. *Journal of Risk and Uncertainty* 32 (2006) 101-130.

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

[48] Varian, Hal. 1992. Microeconomic Analysis, New York: W. W. Norton.

[49] Williams. A. and Lund, A. (1986) "Seat belt use laws and occupant crash protection in the United States." *American Journal of Public Health*, 76: 1438-42.