



**University of
Leicester**

DEPARTMENT OF ECONOMICS

The Behavioral Economics of Insurance

Ali al-Nowaihi, University of Leicester, UK

Sanjit Dhami, University of Leicester, UK

Working Paper No. 10/12

Updated April 2010

The Behavioral Economics of Insurance and the Composite Prelec Probability Weighting Function.

Ali al-Nowaihi* Sanjit Dhami†

28 April 2010

Abstract

We focus on four stylized facts of behavior under risk. Decision makers: (1) Overweight low probabilities and underweight high probabilities. (2) Ignore events of extremely low probability and treat extremely high probability events as certain. (3) Buy inadequate insurance for very low probability events. (4) Keeping the expected loss fixed, there is a probability below which the take-up of insurance drops dramatically. Expected utility (EU) fails on 1-4. Existing models of rank dependent utility (RDU) and cumulative prospect theory (CP) satisfy 1 but fail on 2, 3, 4. We propose a new class of axiomatically-founded probability weighting functions, the composite Prelec weighting functions (CPF) that simultaneously account for 1 and 2. When CPF are combined with RDU and CP we get respectively, *composite rank dependent utility* (CRDU) and *composite cumulative prospect theory* (CCP). Both CRDU and CCP are able to successfully explain 1-4. CCP is, however, more satisfactory than CRDU because it incorporates the empirically robust phenomena of reference dependence and loss aversion.

Keywords: Decision making under risk, Insurance, Composite Prelec probability weighting functions, Composite rank dependent utility theory, Composite cumulative prospect theory, power invariance, local power invariance.

JEL Classification: C60(General: Mathematical methods and programming), D81(Criteria for decision making under risk and uncertainty).

*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.

“... people may refuse to worry about losses whose probability is below some threshold. Probabilities below the threshold are treated as zero.” Kunreuther et al. (1978, p. 182).

“Obviously in some sense it is right that he or she be less aware of low probability events, other things being equal; but it does appear from the data that the sensitivity goes down too rapidly as the probability decreases.” Kenneth Arrow in Kunreuther et al. (1978, p. viii).

“An important form of simplification involves the discarding of extremely unlikely outcomes.” Kahneman and Tversky (1979, p. 275).

“Individuals seem to buy insurance only when the probability of risk is above a threshold ...” Camerer and Kunreuther et al. (1989, p. 570).

1. Introduction

The insurance industry is of tremendous economic importance. The total global gross insurance premiums for 2008 were 4.27 trillion dollars, which accounted for 6.18% of global GDP (Plunkett, 2010). The study of insurance is crucial in almost all branches of economics. Yet, despite impressive progress, we show that existing theoretical models are unable to explain the stylized facts on the take-up of insurance for *low probability events*. The main motivation for the paper is to provide a theory that explains the stylized facts on insurance for events of all probabilities, particularly those occurring with low probabilities.

We highlight the following four robust empirical stylized facts, S1-S4, about the behavior of decision makers under risk.

- S1. Decision makers overweight low probabilities but underweight high probabilities.
- S2. Decision makers ignore very low probability events and treat very high probability events as certain.
- S3. Decision makers buy inadequate insurance against low probability events.
- S4. Whether premiums are actuarially fair, unfair or subsidized, there is a probability below which the take-up of insurance drops dramatically, as the probability of the loss decreases and the loss increases, keeping the expected loss constant.

We discuss these stylized facts at greater length, below. For the moment, note that S1 and S2 are generic stylized facts about human behavior under risk.¹ However, S3, S4 apply specifically in an insurance context.²

¹See, for instance, Kahneman and Tversky (1979), Kahneman and Tversky (2000), Starmer (2000).

²See Kunreuther (1978). S3 and S4 are not the only known behavioral attitudes towards insurance. Individuals might also be influenced by framing effects and exhibit the conjunction fallacy when they view damages broken into their subparts; see Camerer and Kunreuther (1989). These lie outside the scope of this paper.

The most popular decision theory in economics, *expected utility theory* (EU), predicts too much insurance and it violates all of S1-S4. For example, a well known result under EU is that a risk-averse decision maker will buy full insurance if the premium is actuarially fair. By contrast, Kunreuther et al. (1978, p.169) found that only 20% of experimental subjects buy insurance, at the actuarially fair premium, if the probability of the loss is 0.001. We shall consider the insurance implications of EU in more detail in Section 2.3.

Rank dependent utility theory (RDU) and *cumulative prospect theory* (CP) were both developed to give a better explanation of behavior under risk. RDU and CP, unlike EU, incorporate *probability weighting functions*³ that incorporate stylized fact S1. An example is the axiomatically-founded Prelec (1998) probability weighting function that is consistent with S1 (but, crucially, not S2). We show that this feature leads to the excessive take-up of insurance. In this sense, RDU and CP actually do worse than EU. The main aim of our paper is to propose a theoretical solution that enables RDU and CP to explain the stylized facts S1-S4. However, it is not possible to rescue EU.

To cope with the stylized fact S2 (and other phenomena) Kahneman and Tversky (1979) proposed an *editing phase* where decision makers decide which events to ignore and which to treat as certain. This is followed by an *evaluation phase* in which the decision maker chooses from among the *psychologically-cleaned* lotteries from the first stage.

Inspired by Kahneman and Tversky's (1979) idea,⁴ we propose a new class of probability weighting functions which are simultaneously able to account for S1 and S2. We call these *composite Prelec probability weighting functions* (CPF) because they are composed of several segments of Prelec (1998) probability weighting functions, each appropriate for the relevant probability range. The CPF's are parsimonious, flexible, compatible with the empirical evidence and are axiomatically-founded (see Appendix 2, Proposition 14, for an axiomatic derivation of CPF).

We use the term *composite rank dependent utility theory* (CRDU) to refer to (otherwise standard) RDU when combined with a CPF. Similarly, we use the term *composite prospect theory* (CCP) to refer to (otherwise standard) CP when combined with a CPF. We show how CCP and CRDU can explain all the stylized facts S1-S4. We know of no other decision theory under risk that can replicate this performance. Furthermore, we argue that CCP is likely to be superior to CRDU because it also incorporates reference dependence and loss aversion which are not only psychologically salient but are also robust empirical findings.⁵ We conjecture that CCP can be fruitfully applied to all situations of risk, especially those

³By a *probability weighting function* we mean a strictly increasing function $w(p) : [0, 1] \xrightarrow{onto} [0, 1]$. $w(p)$ is the subjective probability assigned by an individual to an objective probability, p .

⁴Incidentally Kahneman and Tversky (1979) is the second most cited paper in *Econometrica* and in all of economics. We are grateful to Peter Wakker for pointing this out.

⁵See, for instance, Kahneman and Tversky (2000), Starmer (2000) and Dhami and al-Nowaihi (2007, 2010b).

where the probability of an event is very low.⁶

The structure of the paper is as follows. In Section 2 we discuss, *heuristically*, human behavior for low probability events, and the relative failure of EU, RDU and CP to explain observed insurance behavior. Section 3 *formally* describes the basic insurance model. Section 4 introduces some essentials of non-linear probability weighting functions that are required to apply RDU and CP. Section 5 describes the Prelec probability weighting function (Prelec, 1998), which plays a fundamental role in this paper. In Sections 6 and 7 we formally derive our results on insurance behavior under RDU and CP, respectively. Section 8 introduces our proposed new probability weighting functions, which we call composite Prelec probability weighting functions (CPF). Section 8 then combines CPF with, respectively, RDU and CP to form composite rank dependent theory (CRDU) and composite prospect theory (CCP). Section 9 shows that CRDU and CCP successfully explain S1-S4. Section 10 concludes. Proofs are relegated to Appendix 1. Appendix 2 contains an axiomatic derivation of the CPF (Proposition 14) and some useful information on the Prelec function (Propositions 12 and 13).

2. The insurance problem

2.1. Insurance for low probability events (Stylized facts S3, S4)

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 (*stylized fact S3*). This was a major study, with 135 expert contributors, involving samples of thousands, survey data, econometric analysis and experimental evidence. All three methodologies gave rise to the same conclusion, see Kunreuther et al. (1978).⁷

EU predicts that a decision maker facing an actuarially fair premium will buy full insurance for all probabilities, however small. Kunreuther et al. (1978, chapter 7) report the following experimental results. They 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 (*stylized fact S4*). These results were robust to changes in subject population, experimental format and order of presentation, presenting the risks separately or simultaneously, bundling the risks, compounding over time and introducing ‘no claims bonuses’.

⁶See al-Nowaihi and Dhimi (2010) for further details.

⁷In the foreword, Arrow (Kunreuther et al., 1978, p. vii) writes: “The following study is path breaking in opening up a new field of inquiry, the large scale field study of risk-taking behavior.”

Remarkably, the lack of interest in buying insurance arose despite active government attempts to (i) provide subsidy to overcome transaction costs, (ii) reduce premiums below their actuarially fair rates, (iii) provide reinsurance for firms and (iv) provide relevant information. Hence, one can safely rule out these factors as contributing to the low take-up of insurance. 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.”

A skeptical reader might question this evidence on the grounds that potential buyers of insurance could have had limited awareness of the losses or that they might be subjected to moral hazard (in the expectation of federal aid). Both explanations are rejected by the data in Kunreuther et al. (1978). Furthermore, there is no evidence of procrastination (arising from say, hyperbolic discounting) in the Kunreuther et al. (1978) data.

2.2. Other examples of individual response to low probability events

In diverse contexts, people ignore low probability events that could, in principle impose huge losses to them (stylized fact S2). Since many of these losses are self-imposed, because of individual actions, people are choosing not to self-insure. It is beyond the scope of this paper to go through the relevant evidence, but we make some suggestive remarks here.⁸

People were reluctant to use seat belts prior to their mandatory use despite publicly available evidence that seat belts save lives. Prior to 1985, only 10-20% of motorists wore seat belts voluntarily, hence, denying themselves self-insurance; see Williams and Lund (1986). Even as evidence accumulated about the dangers of breast cancer (low probability event) women took up the offer of breast cancer examination, only sparingly.⁹

Bar-Ilan and Sacerdote (2004) 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, however, the consequences are self inflicted and potentially have infinite costs.

A user of mobile phones, while driving, faces potentially infinite costs (e.g. loss of one’s and/or the family’s life) with low probability, in the event of an accident. Survey evidence in UK indicates that up to 40% of individuals drive and talk on mobile phones; see, Royal Society for the Prevention of Accidents (2005). Pöystia et al. (2004) report that two thirds of Finnish drivers and 85% of American drivers use their phone while driving. Mobile phone usage, while driving, increases the risk of an accident by two to

⁸See Dhami and al-Nowaihi (2010a) for further examples.

⁹In the US, this changed after the greatly publicised events of the mastectomies of Betty Ford and Happy Rockefeller; see Kunreuther et al. (1978, p. xiii and p. 13-14).

six fold. Hands-free equipment, although now obligatory in many countries, seems not to offer essential safety advantage over hand-held units. A natural explanation is that the individuals simply ignore or substantially underweight the low probability of an accident.

2.3. The failure of expected utility theory (EU) to explain the stylized facts

It is a standard theorem of EU that people will insure fully if, and only if, they face actuarially fair premiums. Since insurance firms have to at least cover their costs, market premiums have to be above the actuarially fair ones. Thus, EU provides a completely rational explanation of the widely observed phenomenon of under-insurance. This has the policy implication that if full-insurance is deemed necessary (because of strong externalities for example), then it has to be encouraged through subsidy or stipulated by law. However, EU is unable to explain several stylized facts from insurance, as we now outline.¹⁰

First, Kunreuther et al. (1978, ch.7) found that only 20% of experimental subjects insure, at the actuarially fair premium, if the probability of the loss is 0.001. Second, EU cannot explain why many people simultaneously gamble and insure. The size of the gambling/insurance industries makes it difficult to dismiss such behavior as quirky. Third, EU predicts that a risk averse decision maker always buys some insurance, even when premiums are unfair. However, many people simply do not buy any insurance, even when it is available, especially for low probability events. Fourth, when faced with an actuarially unfair premium, EU predicts that a decision maker, who is indifferent between full-insurance and not insuring, would strictly prefer *probabilistic insurance*¹¹ to either. This is contradicted by the experimental evidence (Kahneman and Tversky, 1979: 269-271).

What accounts for the low take-up of insurance for low probability events? There is some evidence of a bimodal perception of risks that could offer a potential explanation.¹² 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, the former are likely to ignore insurance while the latter might buy it for low probability but highly salient events. We focus here on the former set of individuals.

¹⁰There are also well known problems with EU in non-insurance contexts; see Kahneman and Tversky (2000), Starmer (2000) and Dhami and al-Nowaihi (2007, 2010b).

¹¹For example, according to EU, an individual who is indifferent between full insurance and not insuring at all should strictly prefer being covered on (say) even days (but not odd days) to either.

¹²See Camerer and Kunreuther (1989) and for the evidence, see McClelland et al (1993) and Schade et al (2001).

2.4. The failures of alternative theories to explain the stylized facts

The difficulties arising from the use of EU have motivated a number of alternatives. The most important of these are *rank dependent utility theory* (RDU), see Quiggin (1982, 1993), and *cumulative prospect theory* (CP), see Tversky and Kahneman (1992).¹³ These two theories offer several improvements over EU, e.g., CP can explain the anomalous result arising from probabilistic insurance that EU cannot.¹⁴

Unlike EU, both RDU and CP use *probability weighting functions*, $w(p)$, to overweight low probabilities and underweight high probabilities (stylized fact S1). One such $w(p)$ function, that is consistent with much of the evidence on non-extreme probability events, and has axiomatic foundations, is the Prelec (1998) function plotted in Figure 2.1 below.¹⁵

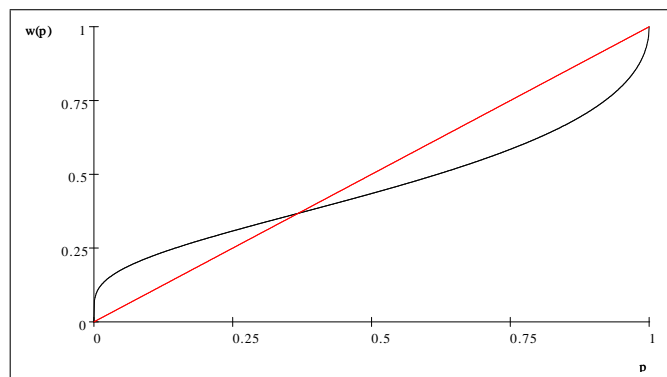


Figure 2.1: The Prelec (1998) function, $w(p) = e^{-(-\ln p)^{\frac{1}{2}}}$.

Remark 1 (*Infinite overweighting of infinitesimal probabilities*): Several weighting functions, in addition to Prelec’s (1998), have been proposed; we discuss some below. However, they all have the feature that the decision maker, (1) *infinitely overweight infinitesimal probabilities* in the sense that the ratio $w(p)/p$ goes to infinity as p goes to zero, and (2) *infinitely underweight near-one probabilities* in the sense that the ratio $[1 - w(p)]/[1 - p]$ goes to infinity as p goes to 1. We call the set of these functions as *standard probability weighting functions*. They underpin all models of RDU and CP and violate stylized fact S2.

¹³Mark Machina (2008) has recently argued that “... the Rank Dependent form has emerged as the most widely adopted model in both theoretical and applied analyses.” However, CP, by adding the notions of reference dependence, loss aversion and separate attitudes to risk in the domain of gains and losses, can explain everything that RDU can do and, in addition, explain phenomena of major economic importance that RDU cannot.

¹⁴For other specific improvements offered by RDU and CP over EU, see Starmer (2000).

¹⁵See Definition 5, below.

We make three contributions, C1-C3, in this paper. The first contribution, C1, is listed immediately below, followed by contributions C2 and C3 in the next subsection.

- C1. *Failure of RDU and CP*: Recall that ‘standard probability weighting functions’ infinitely overweight infinitesimal probabilities. Hence, we show that decision makers who use RDU or CP, insure fully against a sufficiently low probability loss even under actuarially unfair premiums and fixed costs of insurance, provided a mild participation constraint is satisfied. *Thus, CP and RDU cannot account for stylized facts S3 and S4.* This is a big blow to RDU and CP, because both were proposed precisely to overcome the empirical refutations of EU in the domain of choice under risk.

2.5. Towards an explanation of the stylized facts from insurance

We now describe how RDU and CP might be *modified* to explain the stylized facts S1-S4. We begin with the insight of Kahneman and Tversky (1979) in *prospect theory* (PT), which describes two sequential phases.

1. In the *editing phase*, decision makers choose which improbable events to treat as impossible and which probable events to treat as certain. In the insurance context, decision makers might ignore events below a probability threshold (stylized fact S3). Kahneman and Tversky (1979, pp. 282-83) wrote: “*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.*”
2. In the *decision/evaluation phase*, which follows the editing phase, decision makers apply prospect theory to the psychologically-edited lotteries. Having decided to ignore low probability events in the editing phase, decision makers demand zero insurance for such events in the decision phase, in conformity with stylized fact S3.

We can now describe our remaining two main contributions, C2 and C3.

- C2. *Composite Prelec probability weighting functions (CPF)*: Recall that the Prelec weighting function (see Figure 2.1) captures stylized fact S1 but fails on stylized fact S2. We modify the end points of the Prelec function that enables us to address, simultaneously, S1 and S2. Intuitively, we combine the editing and the evaluation phases of PT. The resulting function, the *composite Prelec probability weighting function* (CPF) is sketched in Figure 2.2. It is *consistent with the empirical evidence*, is *parsimonious* but *flexible* and has an *axiomatic foundation* (see Appendix 2, Proposition 14, for an axiomatic derivation of CPF). In Figure 2.2, decision makers heavily underweight

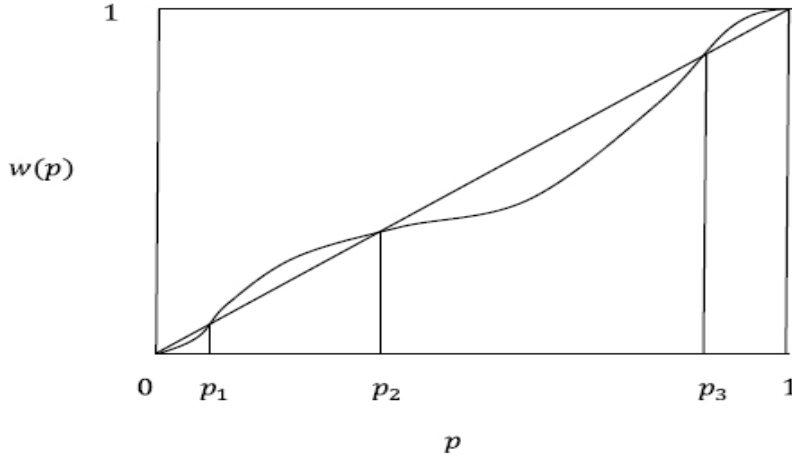


Figure 2.2: A composite Prelec probability weighting function (CPF)

very low probabilities in the range $[0, p_1]$ (by contrast the Prelec weighting function in Figure 2.1, infinitely overweights near-zero probabilities). Akin to Kahneman and Tversky’s (1979) editing phase, decision makers who use the CPF in Figure 2.2 would typically ignore very low probability events by assigning low subjective weights to them (stylized fact S2). Hence, in conformity with the evidence, they do not buy insurance for very low probability events (unless mandatory).¹⁶ Over the probability range $[p_3, 1]$, decision makers overweight the probability, as suggested by the evidence; see, for instance, Kahneman and Tversky’s (1979, p.282-83). In the middle segment, $p \in [p_1, p_3]$, the CPF is identical to the Prelec function, and so addresses stylized fact S1.

C3. *Explanation of S1-S4 using composite prospect theory (CCP: CP+CPF) and composite rank dependent utility (CRDU: RDU+CPF):* Tversky and Kahneman (1992) introduced *cumulative prospect theory* (CP). This replaced Kahneman and Tversky’s (1979) *prospect theory* (PT) in two respects. (i) The psychologically-rich editing phase that determined, among other things, which low probability events to ignore (stylized fact S2) was eliminated.¹⁷ (ii) Cumulative transformations of probability replaced the point transformations under PT.¹⁸ Hence, a decision maker using CP

¹⁶In other contexts, they are unlikely to be dissuaded from low-probability high-punishment crimes, reluctant to wear seat belts (unless mandatory), reluctant to participate in voluntary breast screening programs (unless mandatory) and so on.

¹⁷The reason for this was that the cumulative transformations of probabilities in CP requires a continuous, 1 – 1, and onto probability weighting function, $w(p)$, on $[0, 1]$ such that $w(0) = 0$ and $w(1) = 1$. The editing phase in PT, however, creates discontinuities in the weighting function, which are not permissible under CP.

¹⁸The introduction of cumulative transformation of probabilities was an insight borrowed from RDU.

never chooses stochastically dominated options (unlike PT).

When CP is augmented with CPF (instead of a standard probability weighting function; see remark 1) we refer to it as *composite cumulative prospect theory* (CCP). Analogously, when combined with CPF, RDU is referred to as *composite rank dependent utility*, CRDU. CCP and CRDU are able to address all four stylized facts, S1-S4. In particular, a decision maker who uses CCP and CRDU will not buy insurance against an expected loss of sufficiently low probability; in agreement with the evidence. To quote from Kunreuther et al. (1978, p248) “This brings us to the key finding of our study. The principal reason for a failure of the market is that most individuals do not use insurance as a means of transferring risk from themselves to others. This behavior is caused by people’s refusal to worry about losses whose probability is below some threshold.”

2.6. Are these low probabilities economically relevant?

In section 8.1, we fit two composite Prelec probability weighting functions (CPF) to two separate data sets from Kunreuther et al. (1978). The range of low probabilities that are underweighted by the CPF (the range $(0, p_1]$ in Figure 2.2) are, respectively, $(0, 0.195]$ and $(0, 0.006]$. The upper ends of these intervals involve probabilities that would comfortably seem to accommodate many insurance contexts. The decision maker might, however, still want to buy insurance even if he/she underweights these probabilities. This brings us to the issue of the exact amount of underweighting and the probability of a loss for which insurance is sought. These are some of the issues addressed in our paper.

3. The Model

Suppose that a decision maker can suffer the loss, $L > 0$, with probability $p \in (0, 1)$. She can buy coverage, $C \in [0, L]$, at the cost $rC + f$, where $r \in (0, 1)$ is the *premium rate*, and $f \geq 0$ is a *fixed cost* of buying insurance.¹⁹ We allow departures from the *actuarially fair condition*. We do so in a simple way by setting the insurance premium rate, r , to be

$$r = (1 + \theta)p. \tag{3.1}$$

Thus, $\theta = 0$ corresponds to the actuarially fair condition, $\theta > 0$ to the actuarially unfair condition and $\theta < 0$ to the actuarially ‘over-fair’ condition. Hence, the decision maker’s

See, for example, Quiggin (1993) or Starmer (2000) for arguments for preferring the cumulative transformation of probability to the point transformation.

¹⁹The fixed costs include various transactions costs of buying insurance that are not reflected in the insurance premium. For instance, the costs of information acquisition about alternative insurance policies, the opportunity costs of one’s time spent in buying insurance and so on.

wealth is:

$$\begin{cases} W - rC - f & \text{with probability } p \\ W - rC - f - L + C \leq W - rC - f & \text{with probability } 1 - p \end{cases}$$

Let $U_I(C)$ be the utility of the decision maker if she decides to buy an amount of coverage, $C > 0$. Let C^* be the optimal level of coverage. Denote by U_{NI} the utility of the decision maker from not buying any insurance. Then, a decision maker who buys coverage C^* , satisfies the *participation constraint*,

$$U_{NI} \leq U_I(C^*).$$

3.1. Some observations on limiting processes

In this paper, we shall make heavy use of limiting arguments. The decision maker faces a loss, L , with probability, p . Hence, the expected value of the loss is $\bar{L} = pL$. We shall take the limit $p \rightarrow 0$, *keeping \bar{L} fixed*. It follows that $L \rightarrow \infty$. One can object by saying that losses can never be infinite.²⁰ At this level of generality, one could argue against the use of asymptotic theory in statistics on the grounds that samples can never be infinite. One could also argue against using infinite horizon models, against infinitely divisible quantities and so on.²¹

A further example may help. The early literature that studied convergence of Cournot competition to perfect competition assumed that market size remained fixed while the number of producers tended to infinity.²² A consequence was that the size of each producer had to approach zero. Further consequences were that the average cost curve had to be positive sloping and that any transaction cost, however small, would block trade. These consequences were regarded as unsatisfactory. Therefore, the subsequent literature considered producers of fixed size with U-shaped average cost curves, possibly with fixed costs. As the number of firms tended to infinity, the total size of the market had to also tend to infinity. Although no market can, in fact, be infinite, the later literature was regarded as more satisfactory.²³

²⁰Actually, loss of life, which can be associated with many kinds of low probability disasters is arguably an infinite loss. Loss of honour or severe injustice (in many societies), or suffering a serious lifelong handicap could also approximate infinite losses.

²¹Continuous-time trading and an exogenous process for the underlying asset are fundamental assumptions of the Black-Scholes-Merton model, one of the fundamental models in finance. However, any transaction costs, no matter how small, would block continuous time trading. An exogenous process for the underlying asset (unaffected by trade in the financial derivative) implies an infinite quantity traded in the underlying asset. Similarly, in physics, any particle following a Brownian motion must cover an infinite distance in any time period, no matter how brief, which in itself is not a physically acceptable result. But the assumption of Brownian motion is of immense utility, as it facilitates derivation of refutable empirically important consequences.

²²See, for example, Frank (1965) and Ruffin (1971).

²³See, for example, Novshek (1980).

More specifically, one can object by saying that we are considering the *wrong* limiting argument. In particular, that we should take the limit $p \rightarrow 0$, *keeping L fixed*. It then follows that $\bar{L} \rightarrow 0$. However, in fact, the first type of limiting argument ($p \rightarrow 0$, keeping \bar{L} fixed) is the relevant one in our case. The limiting process we consider is the theoretical analogue of the experimental treatments of Kunreuther et al. (1978, chapter 7). They 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.

4. Non-linear transformation of probabilities

The main alternatives to expected utility (EU) under risk, i.e., rank dependent utility (RDU) and cumulative prospect theory (CP), introduce non-linear transformation of the cumulative probability distribution. In this section, we introduce the concept of a probability weighting function and some other concepts which are crucial for the paper.

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

Proposition 1 : A probability weighting function, $w(p)$, 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.

Definition 2 : The function, $w(p)$, (a) *infinitely-overweights infinitesimal probabilities*, if $\lim_{p \rightarrow 0} \frac{w(p)}{p} = \infty$, and (b) *infinitely-underweights near-one probabilities*, if $\lim_{p \rightarrow 1} \frac{1-w(p)}{1-p} = \infty$.

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

Data from experimental and field evidence typically suggest that decision makers exhibit an inverse S-shaped probability weighting function over outcomes. See Figure 5.1 for an example. Tversky and Kahneman (1992) propose the following probability weighting function, where the lower bound on σ comes from Rieger and Wang (2006).

Definition 4 : The Tversky and Kahneman probability weighting function is given by

$$w(p) = \frac{p^\sigma}{[p^\sigma + (1-p)^\sigma]^{\frac{1}{\sigma}}}, \quad 0.5 \leq \sigma < 1, \quad 0 \leq p \leq 1. \quad (4.1)$$

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

Remark 2 (Standard probability weighting functions): A large number of other probability weighting functions have been proposed, e.g., those by Gonzalez 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.

5. Prelec's probability weighting function

The Prelec (1998) probability weighting function has the following merits: parsimony, consistency with much of the available empirical evidence (at least away from the endpoints of the interval $[0, 1]$) and an axiomatic foundation.

Definition 5 (Prelec, 1998): By the Prelec function we mean the probability weighting function $w(p) : [0, 1] \rightarrow [0, 1]$ given by

$$w(0) = 0, w(1) = 1, \tag{5.1}$$

$$w(p) = e^{-\beta(-\ln p)^\alpha}, 0 < p \leq 1, \alpha > 0, \beta > 0. \tag{5.2}$$

Proposition 3 : The Prelec function (Definition 5) is a probability weighting function in the sense of Definition 1.

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 $w(p) = e^{-(-\ln p)^{\frac{1}{2}}}$ ($\alpha = \frac{1}{2}, \beta = 1$), and sketched in Figure 5.1.

The converse holds if $\alpha > 1$, in which case, the Prelec function is strictly convex for low probabilities but strictly concave for high probabilities, i.e., it is *S-shaped*. An example is $w(p) = e^{-(-\ln p)^2}$ ($\alpha = 2, \beta = 1$), sketched in Figure 5.2 as the light, lower, curve (the straight line in Figure 5.2 is the 45° line).

Between the region of strict convexity ($w'' > 0$) and the region of strict concavity ($w'' < 0$), for each of the cases in Figures 5.1 and 5.2, there is a point of inflexion ($w''(\tilde{p}) = 0$). The parameter β in the Prelec function controls the location of the inflexion point relative to the 45° line. Thus, for $\beta = 1$, this point of inflexion is at $p = e^{-1}$ and lies on the 45° line, as in Figures 5.1 and 5.2 (light curve), above. However, if $\beta < 1$, then the point of inflexion lies above the 45° line, as in $w(p) = e^{-\frac{1}{2}(-\ln p)^2}$ ($\alpha = 2, \beta = \frac{1}{2}$), sketched as

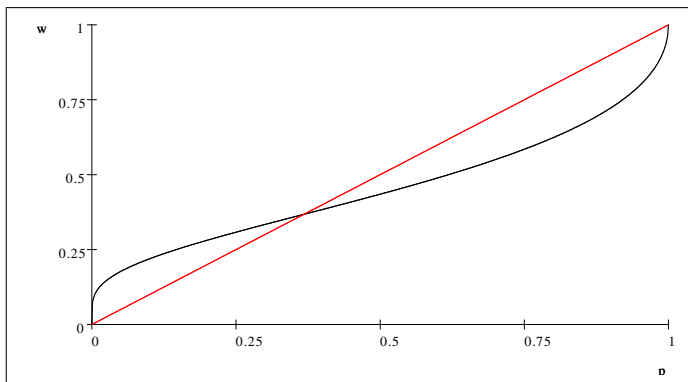


Figure 5.1: A plot of the Prelec (1998) function, $w(p) = e^{-(-\ln p)^{\frac{1}{2}}}$.

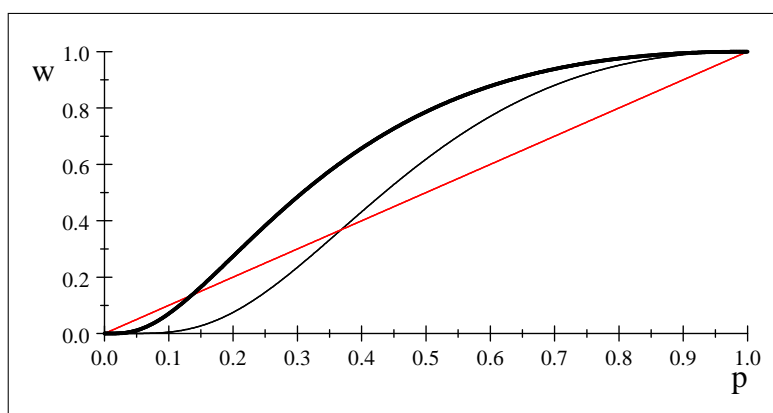


Figure 5.2: A plot of $w(p) = e^{-\frac{1}{2}(-\ln p)^2}$ and $w(p) = e^{-(-\ln p)^2}$.

the thicker, upper, curve in Figure 5.2. In this case, the fixed point, $w(p^*) = p^*$, is at $p^* \simeq 0.14$ but the point of inflexion, $w''(\tilde{p}) = 0$, is at $\tilde{p} \simeq 0.20$.

The full set of possibilities for the Prelec function is established by Propositions 12 and 13 of Appendix 2.

In Figure 5.1 (and first row in Table 2 of Appendix 2), where $\alpha < 1$, note that the slope of $w(p)$ becomes very steep near $p = 0$. By contrast, in figure 5.2 (and last row in Table 2 of Appendix 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.

Proposition 4 : (a) For $\alpha < 1$ the Prelec function (Definition 5): (i) infinitely-overweights infinitesimal probabilities, i.e., $\lim_{p \rightarrow 0} \frac{w(p)}{p} = \infty$, and (ii) infinitely underweights near-one probabilities, i.e., $\lim_{p \rightarrow 1} \frac{1-w(p)}{1-p} = \infty$ (Prelec, 1998, p505); see Definition 2 and Figure 5.1.

(b) For $\alpha > 1$ the Prelec function: (i) zero-underweights infinitesimal probabilities, i.e., $\lim_{p \rightarrow 0} \frac{w(p)}{p} = 0$, and (ii) zero-overweights near-one probabilities, i.e., $\lim_{p \rightarrow 1} \frac{1-w(p)}{1-p} = 0$; see Defi-

tion 3 and figure 5.2.

According to Prelec (1998, p.505), the infinite limits in Proposition 4a capture the qualitative change as we move from certainty to probability and from impossibility to improbability. On the other hand, they contradict stylized fact S2, 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). In sections 6 and 7, below, we show that this leads to people fully insuring against all losses of sufficiently low probability, even with actuarially unfair premiums and fixed costs of insurance. This is contrary to the evidence, e.g., that in Kunreuther et al. (1978). These specific problems are avoided for $\alpha > 1$. However, for $\alpha > 1$, the Prelec function is S-shaped, see Proposition 13d and Figure 5.2. This is in conflict with the empirical evidence, which indicates an inverse-S shape for probabilities bounded away from the end points of the interval $[0, 1]$ (see stylized fact S1). Hence, the two cases, $\alpha < 1$ or $\alpha > 1$, by themselves, are unable to simultaneously address stylized facts S1 and S2.

6. Rank dependent utility (RDU) and insurance

We now model the behavior of an individual using *rank dependent utility theory* (RDU). RDU may be regarded as a conservative extension of *expected utility theory* (EU) to the case where probabilities are transformed in a non-linear manner using decision weights.

To illustrate RDU, consider a decision maker with utility of wealth, u , and a probability weighting function, $w(p)$, where u is strictly concave, differentiable, and strictly increasing, i.e., $u' > 0$. In addition to these standard assumptions, we shall assume that u' is bounded above²⁴ by (say) u'_{\max} . Suppose that the decision maker faces the lottery $(x_1, p_1; x_2, p_2)$, i.e., the wealth level, x_1 , occurs with probability p_1 and the wealth level, x_2 , occurs with probability p_2 , $x_1 < x_2$, $0 \leq p_i \leq 1$, $p_1 + p_2 = 1$. A decision maker using RDU evaluates the utility from the lottery $(x_1, p_1; x_2, p_2)$, as follows:

$$U(x_1, p_1; x_2, p_2) = [1 - w(1 - p_1)] u(x_1) + w(p_2) u(x_2). \quad (6.1)$$

For $w(p) = p$, RDU reduces to standard expected utility theory.²⁵ Note that the *higher* outcome, x_2 , receives weight $\pi_2 = w(p_2)$, while the *lower* outcome, x_1 , receives weight

²⁴The boundedness of u' is needed for part (b) of Proposition 5. This seems feasible on empirical grounds, since people do undertake activities with a non-zero probability of complete ruin, e.g., using the road, undertaking dangerous sports, etc. However, the boundedness of u' excludes such tractable utility functions as $\ln x$ and x^γ , $0 < \gamma < 1$. By contrast, the boundedness of u' is not a requirement in CP, as we shall see.

²⁵From Proposition 12, this is the case for $\alpha = \beta = 1$ (see Appendix 2).

$\pi_1 = w(p_1 + p_2) - w(p_2) = w(1) - w(p_2) = 1 - w(p_2) = 1 - w(1 - p_1)$. π_1 and π_2 are the *decision weights* associated with the relevant outcomes²⁶

We now apply this framework to the insurance model outlined in section 3. If the decision maker buys insurance, then the utility under RDU is given by:

$$U_I(C) = w(1 - p)u(W - rC - f) + [1 - w(1 - p)]u(W - rC - f - L + C). \quad (6.2)$$

Since $U_I(C)$ is a continuous function on the non-empty compact interval $[0, L]$, an optimal level of coverage, C^* , exists. For full insurance, $C = L$, (6.2) gives:

$$U_I(L) = u(W - rL - f). \quad (6.3)$$

On the other hand, the decision maker's utility from not buying insurance is:

$$U_{NI} = w(1 - p)u(W) + [1 - w(1 - p)]u(W - L). \quad (6.4)$$

The decision maker must satisfy the following *participation constraint* to buy a level of insurance coverage C^* :

$$U_{NI} \leq U_I(C^*). \quad (6.5)$$

Proposition 5 : *Suppose that the decision maker follows RDU.*

(a) *A sufficient condition for the participation constraint in (6.5) to hold is that fixed costs of insurance, f , be bounded above by $\bar{L}F(p)$, where $\bar{L} = pL$ and*

$$F(p) = \frac{1 - w(1 - p)}{p} - (1 + \theta). \quad (6.6)$$

(b) *If the probability weighting function infinitely-underweights near-one probabilities (Definition 2b) then, for a given expected loss, the decision maker will insure fully for all sufficiently small probabilities. This holds even in the presence of actuarially unfair insurance and fixed costs of insurance.*

(c) *If a probability weighting function zero-overweights near-one probabilities (Definition 3b) then, for a given expected loss, a decision maker will not insure, for all sufficiently small probabilities.*

The intuition behind Proposition 5 is as follows. First, consider part (b). Suppose that the probability weighting function infinitely-underweights near-one probabilities (Definition 2b). This is the case for all the *standard probability weighting functions* and, in

²⁶The framework is easily extended to any finite number of outcomes; see Quiggin (1982, 1993) for the details. We now clarify a common source of confusion. The decision weights, π_i (*not* $w(p_i)$), sum up to one under RDU. This is no longer the case under CP when there are both *gains* and *losses* (defined below). However, if all outcomes under CP are in the domain of gains or all are in the domain of losses, then the decision weights will sum to 1.

particular, for the Tversky-Kahneman probability function (Definition 4 and Proposition 2) and for the Prelec function with $\alpha < 1$ (Definition 5 and Proposition 4(aii)). In this case, and as illustrated in Figure 5.1, the probability weighting function is very steep near 1 (becoming infinitely steep in the limit) and considerably underweights probabilities close to 1.

Now return to (6.1). Recall that $x_1 < x_2$, so that (in the present context) x_1 is total wealth if the loss occurs (with probability, p_1) and x_2 is total wealth if the loss does not occur (with probability $p_2 = 1 - p_1$). As $p_1 \rightarrow 0$, $p_2 \rightarrow 1$ and, hence, as required in Proposition 5(b), $w(p_2)$ underweights p_2 . This reduces the relative salience of the weighted utility of wealth, $w(p_2)u(x_2)$, if the loss does not occur but increases the relative salience of the weighted utility of wealth, $[1 - w(p_2)]u(x_1)$, if the loss does occur. This makes insurance even more attractive under RDU than under EU (which was already too high, given the evidence).

The reverse occurs in case (c). Here the probability weighting function zero-overweights near-one probabilities (Definition 3b). This is the case for the Prelec function with $\alpha > 1$ (Definition 5 and Proposition 4(bii)). In this case, and as illustrated in Figure 5.2, the probability weighting function is very shallow near 1 (with the slope, in fact, approaching zero) and considerably overweights probabilities close to 1. Now return to (6.1). As $p_2 \rightarrow 1$, as required in Proposition 5(c), $w(p_2)$ overweights p_2 . This increases the relative salience of the weighted utility of wealth, $w(p_2)u(x_2)$, if the loss does not occur. But it also reduces the relative salience of the weighted utility of wealth, $[1 - w(p_2)]u(x_1)$, if the loss does occur. This makes insurance against very low probability events unattractive under RDU, in conformity with the evidence.

From Proposition 4(aii) and Proposition 5(b), a decision maker using a Prelec probability weighting function (Definition 5), will fully insure against all losses of sufficiently small probability. It is of interest to get a feel for how restrictive this participation constraint is. Example (1), below, suggests that it is a weak restriction.

Example 1 : *To check the restrictiveness of the participation constraint we use the result in Proposition 5(a), i.e., the sufficient condition $f \leq \bar{L}F(p)$ for the participation constraint to hold. The first row of the following Table gives losses, L (in dollars, say), from 10 to 10,000,000, with corresponding probabilities, p , in row 2, ranging from 0.1 to 0.000,000, 1. Hence, the expected loss in each case is $\bar{L} = 1$ and so the sufficient condition is simply*

$$f \leq \frac{1 - w(1 - p)}{p} - (1 + \theta) \equiv F(p).$$

In row 3 are the corresponding values of $\frac{1 - w(1 - p)}{p}$ for the Prelec function $w(p) = e^{-(-\ln p)^{0.65}}$, where the values $\alpha = 0.65$ and $\beta = 1$ are suggested by Prelec (1998). Row 4 of the table

reports the corresponding value for $F(p)$ for the case of a relatively high profit rate for insurance firms of 100% (i.e., $\theta = 1$) so that $F(p) = \frac{1-w(1-p)}{p} - 2$. Row 5 gives the upper bound on fixed costs as a percentage of expected losses..

L	10^1	10^2	10^3	10^4	10^5	10^6	10^7
p	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}
$\frac{1-w(1-p)}{p}$	2.067 4	4.903 9	11.161	25.088	56.219	125.88	281.83
$F(p)$	0.067 4	2.903 9	9.161	23.088	54.219	123.88	279.83
$\frac{F(p)}{L} \times 100$	6.74	290.3 9	916.1	2308.8	5421.9	12388	27983

We see, from the table, that the upper bound on (i) fixed costs, and (ii) fixed cost as a percentage of the expected loss, is hardly restrictive for low probabilities. Thus, from Proposition 5(a), we see that using RDU in combination with the Prelec weighting function is likely to lead to misleading results, in that it would predict too much insurance.

Proposition 5(c) will enable us to demonstrate that the decision maker will not insure against any loss of sufficiently small probability, in agreement with observation. We further discuss this in section 8, below.

7. Cumulative prospect theory (CP) and insurance

The four central features of *cumulative prospect theory* (CP) all based on strong empirical evidence, are as follows.²⁷ (1) In CP, unlike EU and RDU, the carriers of utility are not levels of wealth, assets or goods, but differences between these and a reference point (*reference dependence*). The reference point is often (but not necessarily) taken to be the status quo. (2) The *utility function* in prospect theory is concave for gains but convex for losses (*declining sensitivity*).²⁸ (3) The disutility of a loss is greater than the utility of a gain of the same magnitude (*loss aversion*). (4) Probabilities are transformed, so that small probabilities are overweighted but high probabilities are underweighted. We now explain these features in greater detail.

Consider a level of wealth y and a reference point r . Then define the transformed variable $x = y - r$ as the wealth relative to the reference point. The utility function in CP is defined over x ; $x > 0$ is a gain while $x < 0$ is a loss.

Definition 6 (Tversky and Kahneman, 1979). *The utility function under CP, $v(x) : (-\infty, \infty) \rightarrow (-\infty, \infty)$, is a continuous, strictly increasing, mapping that satisfies:*

1. $v(0) = 0$ (*reference dependence*)

²⁷See, for instance, Kahneman and Tversky (2000), Starmer (2000).

²⁸This feature of CP can also explain the observation that individuals can simultaneously gamble and insure. CP can also explain the unpopularity of probabilistic insurance (see section 2.3).

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).

Example 2 : The properties in Definition 6 are satisfied by $v(x) = 1 - e^{-x}$ and $v(x) = x^\gamma$, $0 < \gamma < 1$ but not by $v(x) = \ln x$, since $\ln 0$ is not defined. A popular utility function in CP that is consistent with the evidence and is axiomatically founded is the power form:²⁹

$$v(x) = \begin{cases} x^\gamma & \text{if } x \geq 0 \\ -\lambda(-x)^\gamma & \text{if } x < 0 \end{cases}, \quad 0 < \gamma < 1, \lambda > 1. \quad (7.1)$$

7.1. Illustration of CP for the case of a prospect with two outcomes

Consider a decision maker who faces the lottery $(x_1, p_1; x_2, p_2)$, i.e., win x_1 with probability p_1 or x_2 with probability p_2 , $x_1 < x_2$, $0 \leq p_i \leq 1$, $p_1 + p_2 = 1$. The case of most interest to us is that when both x_1 and x_2 are in the domain of losses, i.e., $x_1 < x_2 < 0$. In this case, the decision maker using CP evaluates the lottery $(x_1, p_1; x_2, p_2)$, by forming the *value function*, V , as follows:

$$V(x_1, p_1; x_2, p_2) = w^-(p_1) v(x_1) + [1 - w^-(p_1)] v(x_2). \quad (7.2)$$

Note that here, the *lower* outcome, x_1 , receives weight $\pi_1 = w^-(p_1)$, while the *higher* outcome, x_2 , receives weight $\pi_2 = w^-(p_1 + p_2) - w^-(p_1) = w^-(1) - w^-(p_1) = 1 - w^-(p_1)$.³⁰

7.2. The insurance decision of an individual who uses CP

Consider a decision maker whose behavior is described by CP and who faces the insurance problem described in Section 3. Take the status-quo wealth, W , of the decision maker as the reference point. With probability $1 - p$, her wealth relative to the reference wealth is

$$(W - rC - f) - W = -rC - f < 0. \quad (7.5)$$

²⁹For an axiomatic derivation of this value function, see al-Nowaihi et al. (2008).

³⁰If $0 < x_1 < x_2$, then the value function under CP is:

$$V(x_1, p_1; x_2, p_2) = w^+(p_2) v(x_2) + [1 - w^+(p_2)] v(x_1), \quad (7.3)$$

which is similar to the corresponding formula for RDU in (6.1) except that here x_i are deviations from reference wealth, while in (6.1) they are absolute levels of wealth. However, if x_1 is in the domain of losses while x_2 is in the domain of gains, so that $x_1 < 0 < x_2$, then

$$V(x_1, p_1; x_2, p_2) = w^+(p_2) v(x_2) + w^-(p_1) v(x_1). \quad (7.4)$$

Note that the probability weighting function for losses, w^- , may be different from that for gains w^+ , although, empirically, it appears that $w^- = w^+ = w$; see Prelec (1998). This is the assumption that we shall also make.

With the complementary probability p , her wealth relative to the reference point is

$$(W - rC - f - L + C) - W = -rC - f - (L - C) \leq -rC - f < 0. \quad (7.6)$$

Notice from (7.5), (7.6) that the decision maker is in the domain of loss in both states.

Definition 7 (*Major and Minor loss*): We call $L + f + rC$ the major loss and $f + rC$ the minor loss. These losses occur with respective probabilities p and $1 - p$.

The decision maker's value function under CP when some level of insurance coverage $C \in [0, L]$ is purchased is given by:

$$V_I(C) = w(p)v(-rC - f - (L - C)) + [1 - w(p)]v(-rC - f). \quad (7.7)$$

Since $V_I(C)$ is a continuous function on the non-empty compact interval $[0, L]$, an optimal level of coverage, C^* , exists. For full insurance, $C = L$, (7.7) gives:

$$V_I(L) = v(-rL - f). \quad (7.8)$$

On the other hand, if the decision maker does not buy any insurance coverage (i.e., $C = 0$), and so also does not incur the fixed cost ($f = 0$), the value function is (recall that $v(0) = 0$):

$$V_{NI} = w(p)v(-L). \quad (7.9)$$

For the decision maker to buy insurance coverage $C^* \in [0, L]$, the following participation constraint (the analogue here of (6.5)) must be satisfied:

$$V_{NI} \leq V_I(C^*). \quad (7.10)$$

Proposition 6 : Suppose that a decision maker uses CP, then the following hold.

(a) There is a corner solution in the following sense. A decision maker will either choose to insure fully against any loss, i.e., $C^* = L$, or choose zero coverage, i.e., $C^* = 0$, depending on the satisfaction of the participation constraint.³¹ This holds even in the presence of actuarially unfair premiums and a fixed cost of insurance.

(b) For Prelec's probability weighting function (Definition 5), with $\alpha < 1$, for the value function (7.1) and for a given expected loss, the participation constraint (7.10) is satisfied for all sufficiently small probabilities. A sufficient condition for the satisfaction of the participation constraint is that $f < \bar{L}F(p)$ where

$$F(p) = \frac{e^{-\frac{\beta}{\gamma}(-\ln p)^\alpha}}{p} - (1 + \theta), \quad 0 < \alpha < 1, \beta > 0, \gamma > 0. \quad (7.11)$$

(c) If a probability weighting function zero-underweights infinitesimal probabilities (Definition 3a) then, for a given expected loss, a decision maker will not insure against any loss of sufficiently small probability.

³¹The answer, therefore, depends on the parameters of the model and so, we require simulations. See Example 3, below.

The reason that part (a) holds is very simple. Since the value function is strictly convex for losses, a decision maker will always insure fully, if he insures at all. This is, of course, at variance with the evidence from partial coverage.

The intuition behind part (b) is as follows. For $\alpha < 1$, the Prelec function infinitely-overweights infinitesimal probabilities (Definitions 2a and 5 and Proposition 4(ai)). In this case, and as illustrated in Figure 5.1, the probability weighting function is very steep near 0 (becoming infinitely steep in the limit) and considerably overweights probabilities close to 0. Now return to (7.2). Recall that $x_1 < x_2 < 0$ in this case. Hence, x_1 is the difference between wealth if the major loss occurs (with probability, p_1) and reference wealth. On the other hand, x_2 is the difference between wealth if the minor loss occurs (with probability $p_2 = 1 - p_1$) and reference wealth. As $p_1 \rightarrow 0$, and as required by Proposition 6(b), $w(p_1)$ increasingly overweights p_1 . This increases the relative salience of the weighted major loss, $w^-(p_1)v(x_1)$ but reduces the relative salience of the weighted minor loss, $[1 - w^-(p_1)]u(x_1)$. This makes insurance even more attractive under CP than under EU (which was already too high, given the evidence).

The reverse occurs in case (c). Here, the probability weighting function zero-underweights infinitesimal probabilities (Definition 3a). This is the case for the Prelec function with $\alpha > 1$ (Definition 5 and Proposition 4(bii)). In this case, and as illustrated in Figure 5.2, the probability weighting function is very shallow near 0 (the slope, in fact, approaches zero) and considerably underweights probabilities close to 0. Now return to (7.2). As $p_1 \rightarrow 0$, and as required by Proposition 6(c), $w(p_1)$ underweights p_1 . This reduces the relative salience of the weighted major loss, $w(p_1)v(x_1)$ but reduces the relative salience of the weighted minor loss, $[1 - w(p_1)]u(x_2)$. This makes insurance against very low probability events unattractive under CP, in conformity with the evidence.

By Proposition 6(a), a decision maker will insure fully against any loss, provided the participation constraint (7.10) is satisfied, even with fixed costs of insurance and an actuarially unfair premium. By Proposition 6(b), for Prelec's probability weighting function (Definition 5), for the value function (7.1) and for a given expected loss, the participation constraint (7.10) is satisfied for all sufficiently small probabilities. It is of interest to get a feel for how restrictive this participation constraint is. Example (3), below, suggests that it is a weak restriction.

Example 3 : *The the first row of the following Table gives losses (in dollars, say) from 10 to 10,000,000, with corresponding probabilities (row 2) ranging from 0.1 to 0.000,000,1; so that the expected loss in each case is $\bar{L} = 1$. In row 3 are the corresponding values of $\frac{e^{-\frac{\beta}{\gamma}(-\ln p)^\alpha}}{p}$, where the values $\alpha = 0.65$ and $\beta = 1$ are suggested by Prelec (1998) and $\gamma = 0.88$ is suggested by Tversky and Kahneman (1992). Row 4 gives $F(p)$ (see (7.11)) for the high profit rate of 100% ($\theta = 1$) for the insurance firm, so that $F(p) = \frac{e^{-\frac{\beta}{\gamma}(-\ln p)^\alpha}}{p} - 2$.*

Row 5 gives the upper bound on fixed costs as a percentage of expected losses.

L	10^1	10^2	10^3	10^4	10^5	10^6	10^7
p	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}
$\frac{e^{-\frac{1}{0.88}(-\ln p)^{0.65}}}{p}$	1.416 9	4.658 9	18.48	81.342	383.83	1906.3	9852.3
$F(p)$	-0.583 1	2.658 9	16.48	79.342	381.83	1904.3	9848.3
$\frac{F(p)}{L} \times 100$	-58.31	265.89	1648	7934.2	38183	190430	984830

It follows that the participation constraint is satisfied, and is hardly restrictive for low probabilities. Thus, from Proposition 6(a),(b), we see that using CP in combination with the Prelec weighting function, is likely to lead to misleading results, because it predicts too much insurance.

Proposition 6(c) will help us reconcile *composite prospect theory* (CCP), which we describe below, with the evidence from the take-up of insurance for low probability events. This will be further discussed in section 8 below.

8. Resolving the insurance problem.

Recall the four stylized facts, S1-S4 in the introduction. Our aim is to introduce *composite Prelec functions* (CPF), which when combined with either RDU or CP, can explain all of the stylized facts S1-S4. We begin, in section 8.1, by providing two numerical examples of CPF that are motivated by the empirical evidence from Kunreuther (1978). This is followed, in subsection 8.2, by a more formal treatment of CPF.

8.1. Two numerical examples of CPF

8.1.1. The urn experiment in Kunreuther (1978)

An obvious solution that simultaneously addresses stylized facts S1 and S2 is to adopt a 3-piece probability weighting function, as in Figure 2.2, above. Figure 8.1, below gives a numerical example of such a CPF, as we now explain.

The CPF in Figure 8.1 is composed of segments from three Prelec functions, and is given by:

$$w(p) = \begin{cases} e^{-0.61266(-\ln p)^2}, & \text{i.e., } \alpha = 2, \beta = 0.61266, & \text{if } 0 \leq p < 0.25, \\ e^{-(\ln p)^{\frac{1}{2}}}, & \text{i.e., } \alpha = 0.5, \beta = 1, & \text{if } 0.25 \leq p \leq 0.75, \\ e^{-6.4808(-\ln p)^2}, & \text{i.e., } \alpha = 2, \beta = 6.4808, & \text{if } 0.75 < p \leq 1. \end{cases} \quad (8.1)$$

The three segments of the CPF in (8.1) are described as follows.

1. For $0 \leq p < 0.25$, the CPF is identical to the S-shaped Prelec function, $e^{-\beta_0(-\ln p)^{\alpha_0}}$, with $\alpha_0 = 2$, $\beta_0 = 0.61266$. β_0 is chosen to make $w(p)$ continuous at $p = 0.25$.

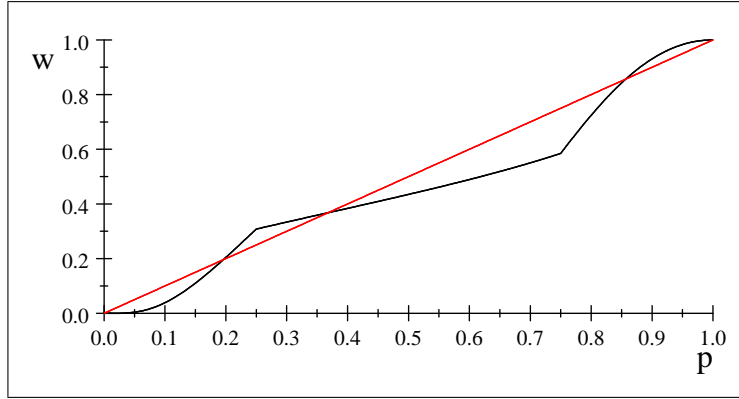


Figure 8.1: The composite Prelec function.

2. For $0.25 \leq p \leq 0.75$, the CPF is identical to the inverse S-shaped Prelec function of Figure 5.1 ($\alpha = 0.5, \beta = 1$).
3. For $0.75 < p \leq 1$, the CPF is identical to the S-shaped Prelec function, $e^{-\beta_1(-\ln p)^{\alpha_1}}$, with $\alpha_1 = 2, \beta_1 = 6.4808$. β_1 is chosen to make $w(p)$ continuous at $p = 0.75$.

Remark 3 (*Fixed points, concavity, convexity*): The CPF in Figure 8.1 has five fixed points, at 0, 0.19549, $e^{-1} = 0.36788$, 0.85701 and 1. It is strictly concave for $0.25 < p < e^{-1}$ and strictly convex for $e^{-1} < p < 0.75$ (a bit hard to see in Figure 8.1, but see our next example).³² The CPF is strictly convex for $0 < p < 0.25$ and strictly concave for $0.75 < p < 1$.

Remark 4 (*Underweighting and overweighting of probabilities*): The CPF in Figure 8.1 overweights ‘low’ probabilities, in the range $0.19549 < p < e^{-1}$ and underweights ‘high’ probabilities, in the range $e^{-1} < p < 0.85701$. These regions, therefore, address stylized fact S1 in the introduction. Furthermore, the CPF underweights ‘very low’ probabilities, in the range $0 < p < 0.19549$ and overweights ‘very high’ probabilities, in the range $0.85701 < p < 1$. For p close to zero, the CPF is nearly flat, thus capturing Arrow’s astute observation: “Obviously in some sense it is right that he or she be less aware of low probability events, other things being equal; but it does appear from the data that the sensitivity goes down too rapidly as the probability decreases.” (Kenneth Arrow in Kunreuther et al., 1978, p. viii). Note that this probability weighting function is also nearly flat near 1. These two segments, i.e., $p \in (0, 0.19549) \cup (0.85701, 1)$ are able to address stylized fact S2.

³²For $\alpha \in (0, \frac{1}{2})$ or $\alpha \in (\frac{1}{2}, 1)$ the concavity/convexity is even milder than for $\alpha = \frac{1}{2}$, with the slope being less steep for $\alpha \in (0, \frac{1}{2})$ but more steep for $\alpha \in (\frac{1}{2}, 1)$. In fact, $w'(p) \rightarrow 0$ as $\alpha \rightarrow 0$ and $w'(p) \rightarrow 1$ as $\alpha \rightarrow 1$.

The parameters in (8.1) have been chosen primarily to clarify the properties mentioned in Remarks 3 and 4. The cutoff points 0.25 and 0.75 in (8.1) and Figure 8.1 were motivated by actual evidence. Kunreuther et al. (1978, chapter 7) report that in one set of their experiments (the “urn” experiments) 80% of subjects (facing actuarially fair premiums) took up insurance against a loss with probability 0.25. But that the take-up of insurance declined when the probability of the loss declined (keeping the expected loss constant). When the probability of the loss reached 0.001, only 20% of the subjects took up actuarially fair insurance. Thus, although Figure 8.1 was chosen primarily for illustration, its qualitative features match the evidence in Kunreuther et al. (1978).

8.1.2. The farm experiments in Kunreuther (1978)

In a second set of experiments, the “farm” experiments, Kunreuther et al. (1978, ch.7) report that the take-up of actuarially fair insurance begins to decline if the probability of the loss (keeping the expected loss constant) goes below 0.05. This is shown in Figure 8.2.

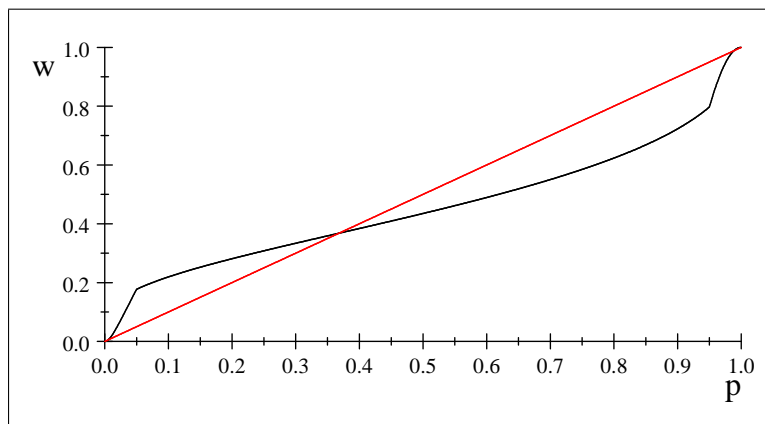


Figure 8.2: The composite Prelec function.

The probability weighting function in Figure 8.2 is composed of segments from three Prelec functions, $w(p) = e^{-\beta(-\ln p)^\alpha}$ and given by:

$$w(p) = \begin{cases} e^{-0.19286(-\ln p)^2}, & \text{i.e., } \alpha = 2, \beta = 0.19286 & \text{if } 0 < p < 0.05 \\ e^{-(-\ln p)^{\frac{1}{2}}}, & \text{i.e., } \alpha = 0.5, \beta = 1 & \text{if } 0.05 \leq p \leq 0.95 \\ e^{-86.081(-\ln p)^2}, & \text{i.e., } \alpha = 2, \beta = 86.081 & \text{if } 0.95 < p \leq 1 \end{cases} \quad (8.2)$$

The three segments of the CPF in (8.2) are described as follows.

1. For $0 \leq p < 0.05$, the CPF is identical to the S-shaped Prelec function, $e^{-\beta_0(-\ln p)^{\alpha_0}}$, with $\alpha_0 = 2$, $\beta_0 = 0.19286$. β_0 is chosen to make $w(p)$ continuous at $p = 0.05$.

2. For $0.05 \leq p \leq 0.95$, the CPF is identical to the inverse S-shaped Prelec function of Figure 5.1 ($\alpha = 0.5, \beta = 1$).
3. For $0.95 < p \leq 1$, the CPF is identical to the S-shaped Prelec function, $e^{-\beta_1(-\ln p)^{\alpha_1}}$, with $\alpha_1 = 2, \beta_1 = 86.081$. β_1 is chosen to make $w(p)$ continuous at $p = 0.95$.

Remark 5 (Fixed points): This CPF has five fixed points: 0, 0.0055993, e^{-1} , 0.98845 and 1. It is strictly concave for $0.05 < p < e^{-1}$ and strictly convex for $e^{-1} < p < 0.95$. It is strictly convex for $0 < p < 0.05$ and strictly concave for $0.95 < p < 1$.

Remark 6 (Underweighting and overweighting of probabilities): The CPF overweightes low probabilities, in the range $0.0055993 < p < e^{-1}$ and underweights high probabilities, in the range $e^{-1} = 0.36788 < p < 0.98845$. This accounts for stylized fact S1. Behavior near $p = 0$, and near $p = 1$, is not obvious from Figure 8.2. So, Figures 8.3 and 8.4, below, respectively magnify the regions near 0 and near 1.

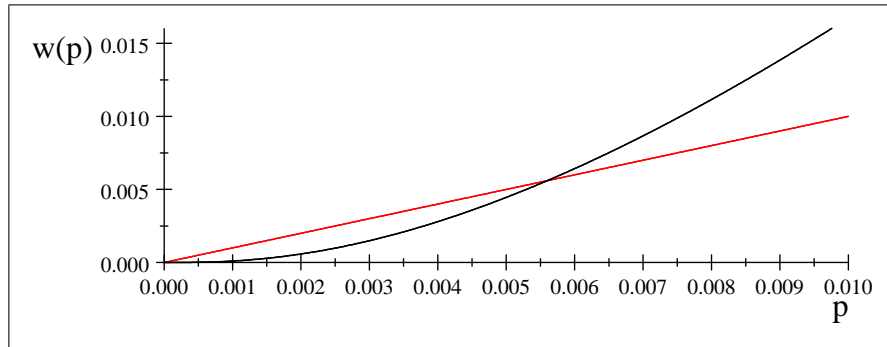


Figure 8.3: Behaviour of Figure 8.2 near 0.

From Figure 8.3, we see that (8.2) underweights very low probabilities, in the range $0 < p < 0.0055993$. For p close to zero, we see that this probability weighting function is nearly flat, thus, again capturing Arrow’s observation “...it does appear from the data that the sensitivity goes down too rapidly as the probability decreases.” From Figure 8.4, we see that (8.2) overweightes very high probabilities, in the range $0.98845 < p < 1$. For p close to one, we see that this probability weighting function is nearly flat.

8.2. A more formal treatment of the CPF

Notice that the upper cutoff points for the *first segment* of the CPF in Figures 8.1 and 8.2 are respectively, at probabilities 0.25 and 0.05. Denote this cutoff point as \underline{p} . Similarly,

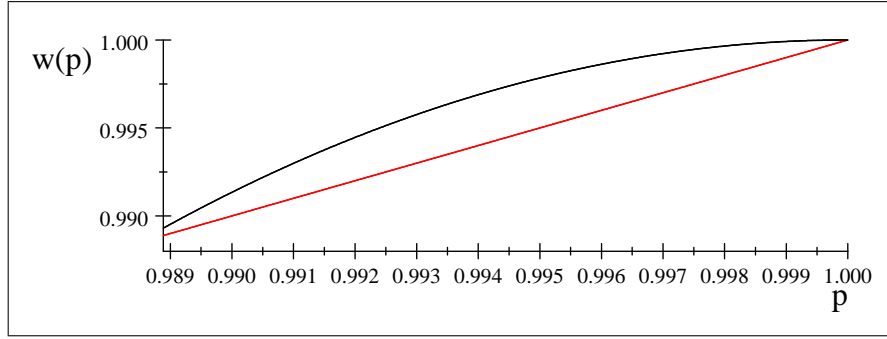


Figure 8.4: Behaviour of Figure 8.2 near 1.

the upper cutoff point for the *second segment* of the CPF in Figures 2.2 and 8.2 are respectively, at probabilities 0.75 and 0.95. Denote this cutoff point as \bar{p} . Now define,

$$\underline{p} = e^{-\left(\frac{\beta}{\beta_0}\right)^{\frac{1}{\alpha_0 - \alpha}}}, \quad \bar{p} = e^{-\left(\frac{\beta}{\beta_1}\right)^{\frac{1}{\alpha_1 - \alpha}}}. \quad (8.3)$$

The probability weighting functions (8.1), (8.2) and their graphs, Figures 8.1-8.4, suggest the following definition.

Definition 8 (*Composite Prelec weighting function, CPF*): By the composite Prelec weighting function we mean the probability weighting function $w : [0, 1] \rightarrow [0, 1]$ given by

$$w(p) = \begin{cases} 0 & \text{if } p = 0 \\ e^{-\beta_0(-\ln p)^{\alpha_0}} & \text{if } 0 < p \leq \underline{p} \\ e^{-\beta(-\ln p)^\alpha} & \text{if } \underline{p} < p \leq \bar{p} \\ e^{-\beta_1(-\ln p)^{\alpha_1}} & \text{if } \bar{p} < p \leq 1 \end{cases} \quad (8.4)$$

where \underline{p} and \bar{p} are given by (8.3) and

$$0 < \alpha < 1, \beta > 0; \alpha_0 > 1, \beta_0 > 0; \alpha_1 > 1, \beta_1 > 0, \beta_0 < 1/\beta^{\frac{\alpha_0 - 1}{1 - \alpha}}, \beta_1 > 1/\beta^{\frac{\alpha_1 - 1}{1 - \alpha}}. \quad (8.5)$$

Proposition 7 : The composite Prelec function (Definition 8) is a probability weighting function in the sense of Definition 1.

The restrictions $\alpha > 0$, $\beta > 0$, $\beta_0 > 0$ and $\beta_1 > 0$, in (8.5), are required by the axiomatic derivations of the Prelec function (see Prelec (1998), Luce (2001) and al-Nowaihi and Dhimi (2006)). The restriction $\beta_0 < 1/\beta^{\frac{\alpha_0 - 1}{1 - \alpha}}$ guarantees that the first segment of the CPF, $e^{-\beta_0(-\ln p)^{\alpha_0}}$, crosses the 45° to the left of \underline{p} and the restriction $\beta_1 > 1/\beta^{\frac{\alpha_1 - 1}{1 - \alpha}}$ guarantees that the third segment of the CPF, $e^{-\beta_1(-\ln p)^{\alpha_1}}$, crosses the 45° degree line to the right of \bar{p} . Together, they imply that the second segment of CPF, $e^{-\beta(-\ln p)^\alpha}$, crosses the 45° between these two limits. It follows that the interval (\underline{p}, \bar{p}) is not empty. The interval

limits are chosen so that the CPF in (8.4) is continuous across them. These observations lead to the following proposition. First, define p_1, p_2, p_3 that correspond to the notation used in Figure 2.2 in the introduction.

$$p_1 = e^{-\left(\frac{1}{\beta_0}\right)^{\frac{1}{\alpha_0-1}}}, p_2 = e^{-\left(\frac{1}{\beta}\right)^{\frac{1}{\alpha-1}}}, p_3 = e^{-\left(\frac{1}{\beta_1}\right)^{\frac{1}{\alpha_1-1}}}. \quad (8.6)$$

Proposition 8 : (a) $p_1 < \underline{p} < p_2 < \bar{p} < p_3$. (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 7, the CPF in (8.4), (8.5) is a probability weighting function in the sense of Definition 1. By Proposition 8, 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 the narrative of Kahneman and Tversky (1979, p. 282-83). Hence, a CPF also accounts for the second stylized fact, S2 in the introduction.

The restrictions $\alpha_0 > 1$ and $\alpha_1 > 1$ in (8.5) ensure that a CPF has the following properties, listed below as Proposition 9, that will help explain human behavior for extremely low probability events; see below.

Proposition 9 : *The CPF (8.4):*

- (a) *zero-underweights infinitesimal probabilities, i.e., $\lim_{p \rightarrow 0} \frac{w(p)}{p} = 0$ (Definition 3a),*
- (b) *zero-overweights near-one probabilities, i.e., $\lim_{p \rightarrow 1} \frac{1-w(p)}{1-p} = 0$ (Definition 3b).*

9. Explaining the stylized facts of insurance

9.1. Composite rank dependent utility theory (CRDU).

Let us use the name *composite rank dependent utility theory* (CRDU) to refer to (otherwise standard) RDU when combined with a composite Prelec probability weighting function (CPF). The following proposition establishes that such a theory would account for empirical facts S3 and S4 in the list at the beginning of this section. Empirical facts S1 and S2 are already satisfied by a CPF, as we saw in section 8.2.

Proposition 10 : *Under composite rank dependent utility theory (CRDU), for a given expected loss, a decision maker will not insure, for all sufficiently small probabilities.*

Thus, unlike EU, RDU or CP, CRDU can account for all four stylized facts S1-S4, listed in the introduction.

9.2. Composite cumulative prospect theory (CCP).

Let us use the name *composite prospect theory* (CCP) to refer to (otherwise standard) CP when combined with a composite Prelec probability weighting function (CPF). The following proposition establishes that such a theory would account for the stylized facts S3 and S4. As shown in section 8.2, S1 and S2 are already satisfied by a CPF.

Proposition 11 : *Under composite prospect theory (CCP), for a given expected loss, a decision maker will not insure, for all sufficiently small probabilities.*

Thus, unlike EU, RDU or CP, CCP can account for all four stylized facts S1-S4. We now provide examples motivated by Kunreuther's (1978) urn experiments; see section 8.1 above.

Example 4 (*Urn experiment*): *Suppose that a decision maker faces a loss, L , of \$200,000 with probability, $p = 0.001$. Insurance is assumed to be actuarially fair, i.e., $r = p$. Under CP, suppose that the decision maker uses the Prelec probability weighting function, $w(p) = e^{-\beta(-\ln p)^\alpha}$ with $\beta = 1$ and $\alpha = 0.50$. Using the experimental values suggested by Kahneman and Tversky (1979), the utility function (7.1) is given by*

$$v(x) = \begin{cases} x^{0.88} & \text{if } x \geq 0 \\ -2.25(-x)^{0.88} & \text{if } x < 0 \end{cases} .$$

For CCP, we take the CPF in (8.1) from Kunreuther's urn experiments. For $p = 0.001$, and for CCP, (8.1) gives $w(0.001) = e^{-0.61266(-\ln 0.001)^2}$. We now check to see if it is optimal for a decision maker under, respectively, CP and CCP, to fully insure, i.e., $C^* = L$. Using (7.10), we need to check, in each case, the following condition that ensures full insurance (we have used the actuarially fair condition $r = p$):

$$w(p)v(-L) \leq v(-pL) \tag{9.1}$$

(a) *Decision maker uses CP: In this case, $w(p) = e^{-\beta(-\ln p)^\alpha}$ with $\beta = 1$ and $\alpha = 0.50$ and (9.1) require that,*

$$e^{-(\ln 0.001)^{0.50}} \left(-2.25 (2 \times 10^5)^{0.88} \right) \leq \left(-2.25 (0.001 \times 2 \times 10^5)^{0.88} \right),$$

$$\Leftrightarrow -7510.2 \leq -238.28,$$

which is true. Hence, a decision maker who uses CP will fully insure. However, Kunreuther's (1978) data shows that only 20% of the decision makers will insure in this case.

(b) *Decision maker uses CCP: In this case, (9.1) and $w(0.001) = e^{-0.61266(-\ln 0.001)^2}$ imply that,*

$$e^{-0.61266(-\ln 0.001)^2} \left(-2.25 (2 \times 10^5)^{0.88} \right) \leq \left(-2.25 (0.001 \times 2 \times 10^5)^{0.88} \right),$$

$$\Leftrightarrow -2.093 \times 10^{-8} \leq -238.28,$$

which is *not true*. Hence, a decision maker using CCP will not insure, which is in conformity with Kunreuther's (1978) data.

Example 5 (*Urn experiment*): Now continue to use the set-up of Example 4. However, let the probability of a loss be $p = 0.25$ (instead of $p = 0.001$). Kunreuther's (1978) data shows that 80% of the experimental subjects took up insurance in this case. For CCP, as in Example 4, we take the CPF in (8.1) from Kunreuther's urn experiments. For $p = 0.25$ the Prelec and CCP functions coincide and $w(p) = e^{-\beta(-\ln p)^\alpha}$. Thus, in each case the full insurance condition $w(p)v(-L) \leq v(-pL)$ in (9.1) is given by,

$$e^{-(-\ln 0.25)^{0.50}} \left(-2.25 (2 \times 10^5)^{0.88} \right) \leq \left(-2.25 (0.25 \times 2 \times 10^5)^{0.88} \right),$$

$$\Leftrightarrow -32044 \leq -30710,$$

which is *true*. Hence, for losses whose probability is bounded well away from the end-points, the insurance predictions of, both, CP and CCP are in conformity with the evidence.

In conjunction, Examples 4 and 5 illustrate how CCP can account well for the evidence for events of all probabilities while CP's predictions for low probability events are incorrect.

10. Conclusion

We argue that a satisfactory theory of insurance should explain the following four stylized facts from the behavior of decision makers under risk. (S1) They overweight low probabilities but underweight high probabilities. (S2) They ignore very low probability events and treat very high probability events as certain. (S3) They buy inadequate insurance against low probability events. (S4) Whether premiums are actuarially fair, unfair or subsidized, and for a fixed expected loss, there is a probability below which the take-up of insurance drops dramatically.

EU fails to account for any of S1-S4, in addition to being subjected to several other problematic aspects in an insurance context that we briefly discuss. *Rank dependent utility* (RDU) and *cumulative prospect theory* (CP) are the two leading alternatives to EU under risk. Both employ non-linear transformations of probability. In addition, CP also incorporates, uniquely among decision theories, the psychologically rich and empirically robust features of *reference dependence* and *loss aversion*, which considerably increase its predictive power relative to RDU.

Both RDU and CP rely on *standard probability weighting functions* to transform objective probabilities. The leading example of a standard probability weighting function is the Prelec (1998) weighting function. All these weighting functions have the feature that they

infinitely overweight infinitesimal probabilities. Decision makers employing these functions are, therefore, eager to insure for very low probability events, contradicting S2, S3, S4. Generally, such decision makers would like to buy even greater insurance than a decision maker who uses EU (whose demand for insurance is already excessive).

We propose a new class of probability weighting functions, the *composite Prelec probability weighting functions* (CPF). A CPF is formed of three segments of *Prelec probability weighting functions*. For probabilities in the middle range, a CPF is identical to a standard Prelec function and satisfies S1, i.e., it is inverse S-shaped. However, it also satisfies stylized fact S2, for probabilities close to the endpoints of the probability interval $[0, 1]$. CPF is parsimonious, flexible, and is axiomatically founded.

When a CPF is combined with otherwise standard RDU we call it as *composite rank dependent utility* (CRDU). Likewise, when a CPF is combined with otherwise standard CP, we call it *composite prospect theory* (CCP). Both, CRDU and CCP are able to account for all four stylized facts S1-S4. On account of the presence of reference dependence and loss aversion in CCP (which are absent in CRDU), we conclude that CCP is the most satisfactory theory of behavior under risk. The implications of CCP are potentially relevant to all economically interesting events under risk, especially those where low probabilities create paradoxes that cannot be resolved with existing theories.³³

11. Appendix 1: Proofs

Proof of Proposition 1: These properties follow almost immediately from Definition 1. ■.

Lemma 1 : Let $w(p)$ be a probability weighting function (Definition 1). Then:

- (a) If $w(p)$ is differentiable in a neighborhood of $p = 0$, then $\lim_{p \rightarrow 0} \frac{w(p)}{p} = \lim_{p \rightarrow 0} w'(p)$.
(b) If $w(p)$ is differentiable in a neighborhood of $p = 1$, then $\lim_{p \rightarrow 1} \frac{1-w(p)}{1-p} = \lim_{p \rightarrow 1} w'(p)$.

Proof of Lemma 1: (a) Let $p \rightarrow 0$. Since w is continuous (Proposition 1c), $w(p) \rightarrow w(0) = 0$ (Proposition 1a). By L'Hospital's rule, $\frac{w(p)}{p} \rightarrow \frac{dw(p)/dp}{dp/dp} = w'(p)$.

(b) Similarly, if $p \rightarrow 1$, then $w(p) \rightarrow w(1) = 1$. By L'Hospital's rule, $\frac{1-w(p)}{1-p} \rightarrow \frac{d[1-w(p)]/dp}{d(1-p)/dp} = w'(p)$. ■.

Proof of Proposition 2: $\frac{w(p)}{p} = \frac{1}{[p^\sigma + (1-p)^\sigma]^{\frac{1}{\sigma}} p^{1-\sigma}} \rightarrow \infty$, as $p \rightarrow 0$.
 $\frac{1-w(p)}{1-p} \rightarrow w'(p) = \left\{ \frac{\sigma}{p} + \frac{1}{p^\sigma + (1-p)^\sigma} \left[\frac{1}{(1-p)^{1-\sigma}} - \frac{1}{p^{1-\sigma}} \right] \right\} w(p) \rightarrow \infty$, as $p \rightarrow 1$. ■.

Proof of Proposition 3: Straightforward from Definition 5. ■.

³³See al-Nowaihi and Dhami (2010) for further details.

Proof of Proposition 4: From (5.2) we get $\ln \frac{w(p)}{p} = \ln w(p) - \ln p = -\beta (-\ln p)^\alpha - \ln p = (-\ln p)^\alpha ((-\ln p)^{1-\alpha} - \beta)$. Hence, if $\alpha < 1$, then $\lim_{p \rightarrow 0} \ln \frac{w(p)}{p} = \infty$ and, hence, $\lim_{p \rightarrow 0} \frac{w(p)}{p} = \infty$. This establishes (ai). On the other hand, if $\alpha > 1$, then $\lim_{p \rightarrow 0} \ln \frac{w(p)}{p} = -\infty$ and, hence, $\lim_{p \rightarrow 0} \frac{w(p)}{p} = 0$. This establishes (bi). From (5.2) we get $w'(p) = \frac{\alpha\beta}{p} (-\ln p)^{\alpha-1} w(p)$. If $\alpha < 1$, then $\lim_{p \rightarrow 1} w'(p) = \infty$. Part (aii) then follows from Lemma 1b. If $\alpha > 1$, then $\lim_{p \rightarrow 1} w'(p) = 0$. Part (bii) then follows from Lemma 1b. ■

Proof of Proposition 5: (a) Consider an expected loss

$$\bar{L} = pL. \quad (11.1)$$

Differentiate (6.2) with respect to C to get

$$\begin{aligned} U'_I(C) &= -rw(1-p)u'(W-rC-f) \\ &\quad + (1-r)(1-w(1-p))u'(W+(1-r)C-f-L). \end{aligned} \quad (11.2)$$

Since u is (strictly) concave, $u' > 0$ and $0 < r < 1$, it follows, from (11.2) that $U'_I(C)$ is a decreasing function of C . Hence,

$$U'_I(L) \leq U'_I(C) \leq U'_I(0) \text{ for all } C \in [0, L]. \quad (11.3)$$

Using (3.1), replace r by $(1+\theta)p$ in (11.2), then divide both sides by p , to get

$$\begin{aligned} \frac{U'_I(C)}{p} &= -(1+\theta)w(1-p)u'(W-(1+\theta)pC-f) \\ &\quad + (1-(1+\theta)p)\frac{1-w(1-p)}{p}u'(W-(1+\theta)pC-f-L+C). \end{aligned} \quad (11.4)$$

For $C = 0$ and $C = L$, (11.4) gives (using (11.1)):

$$\frac{U'_I(0)}{p} = [1-(1+\theta)p]\frac{1-w(1-p)}{p}u'\left(W-f-\frac{\bar{L}}{p}\right) - (1+\theta)w(1-p)u'(W-f), \quad (11.5)$$

$$\frac{U'_I(L)}{p} = \left[\frac{1-w(1-p)}{p} - (1+\theta)\right]u'(W-(1+\theta)\bar{L}-f). \quad (11.6)$$

Since $0 < (1+\theta)p < 1$, $0 < p < 1$, $0 < w(1-p) < 1$, $0 < u' < u'_{\max}$ we get, from (11.5),

$$\frac{U'_I(0)}{p} < \frac{1-w(1-p)}{p}u'_{\max} - (1+\theta)w(1-p)u'(W-f). \quad (11.7)$$

From (11.6) and (6.6) we get,

$$\frac{U'_I(L)}{p} = F(p)u'(W-(1+\theta)\bar{L}-f). \quad (11.8)$$

Since, u' is always positive, from (11.8) we see that

$$U'_I(L) > 0 \Leftrightarrow F(p) > 0. \quad (11.9)$$

From (6.3), (6.4), (11.1), (6.6) and the facts that u is strictly increasing and strictly concave, simple algebra leads to

$$f < \bar{L}F(p) \Rightarrow U_{NI} < U_I(L). \quad (11.10)$$

Thus, if fixed costs are bounded above by $\bar{L}F(p)$, the participation constraint is guaranteed to hold.

Let

$$q = 1 - p. \quad (11.11)$$

(b) Suppose $w(p)$ infinitely-underweights near-one probabilities. Then, from (11.11) and Definition 2b, $\lim_{p \rightarrow 0} \frac{1-w(1-p)}{p} = \lim_{q \rightarrow 1} \frac{1-w(q)}{1-q} = \infty$. Hence, from (6.6), for given expected loss, \bar{L} , we can find a $p_1 \in (0, 1)$ such that, for all $p \in (0, p_1)$, we get $f < \bar{L}F(p)$. From (11.10) it follows that the participation constraint (6.5) is satisfied for all $p \in (0, p_1)$. From $f < \bar{L}F(p)$ we get that $F(p) > 0$ for all $p \in (0, p_1)$. From (11.9) it follows that $U'_I(L) > 0$ for all such p . From (11.3) it follows that $U'_I(C) > 0$ for all such p . Hence, it is optimal for the decision maker to choose as high a coverage as possible, and, so, $C^* = L$, for all $p \in (0, p_1)$, because the participation constraint has already been shown to be satisfied.

(c) Suppose $w(p)$ zero-overweights near-one probabilities. Then, from (11.11) and Definition 3b, $\lim_{p \rightarrow 0} \frac{1-w(1-p)}{p} = \lim_{q \rightarrow 1} \frac{1-w(q)}{1-q} = 0$. Hence, from (11.7), there exists $p_2 \in (0, 1)$ such that for all $p \in (0, p_2)$, $U'_I(0) < 0$. Hence, from (11.3), $U'_I(C) < 0$ for all $C \in [0, L]$. Hence the optimal level of coverage is 0. ■

Proof of Proposition 6: (a) Since v is strictly concave, $-v$ is strictly convex. Hence, from (7.7), it follows that V_I is strictly convex. Since $0 \leq C \leq L$, it follows that $V_I(C)$ is maximized either at $C = 0$ or at $C = L$. Hence, if the participation constraint is satisfied, then the decision maker will fully insure against the loss.

(b) Consider the Prelec function (5.2) and the value function (7.1). Consider an expected loss

$$\bar{L} = pL \quad (11.12)$$

From (5.2), (7.1), (7.8), (7.9), (7.11) and (11.12), simple algebra leads to

$$f < \bar{L}F(p) \Rightarrow V_{NI} < V_I(L). \quad (11.13)$$

From (7.11) and Proposition 4a(i), $\lim_{p \rightarrow 0} F(p) = \infty$. Hence, for given expected loss, \bar{L} , we get $f < \bar{L}F(p)$, for all sufficiently small p . From (11.13) it follows that the participation constraint is satisfied for all such small p .

(c) From (7.8) and (7.9) we get the following

$$\frac{V_I(L) - V_{NI}}{p} = v(L) \frac{w(p)}{p} - v((1 + \theta)\bar{L} + f) \frac{1}{p}, \quad (11.14)$$

$$\lim_{p \rightarrow 0} \frac{V_I(L) - V_{NI}}{p} = v(L) \lim_{p \rightarrow 0} \frac{w(p)}{p} - v((1 + \theta)\bar{L} + f) \lim_{p \rightarrow 0} \frac{1}{p}. \quad (11.15)$$

Suppose $w(p)$ zero-underweights infinitesimal probabilities. Then, from Definition 3a, $\lim_{p \rightarrow 0} \frac{w(p)}{p} = 0$. Hence, the first term in (11.15) goes to 0 as p goes to 0. The second term in (11.15), however, goes to $-\infty$ as p goes to 0. Hence, there exists $p_2 \in (0, 1)$ such that for all $p \in (0, p_2)$, $V_{NI} > V_I(L)$. ■

Proof of Proposition 7: Straightforward from Definitions 1 and 8. ■

Proof of Proposition 8: Follows by direct calculation from (8.4) and (8.5). ■

Proof of Proposition 9: Part (a) follows from part (bi) of Proposition 4, since $\alpha_0 > 1$. Part (b) follows from part (bii) of Proposition 4, since $\alpha_1 > 1$. ■

Proof of Proposition 10: Follows from Propositions 5c and 9b. ■

Proof of Proposition 11: Follows from Propositions 6c and 9a. ■

12. Appendix 2: Axiomatic foundations of the CPF

The main aim of this appendix is to give an axiomatic derivation of the composite Prelec probability weighting function. This is accomplished by Proposition 14 in subsection 12.3, below. This is preceded by subsection 12.2, on Cauchy's famous algebraic functional equations (see, for example, Eichhorn (1978)). These will be of fundamental importance to us. Subsection 12.1, immediately below, gives a fuller treatment of the Prelec (1998) probability weighting function.

12.1. The Prelec probability weighting function

The full set of possibilities for the Prelec function, depending on the parameters α and β , is established by Propositions 12 and 13, below.

Proposition 12 : *For $\alpha = 1$, the Prelec probability weighting function (Definition 5) 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).*

Proof of Proposition 12: Obvious from Definition 5. ■

Lemma 2 : *For $\alpha \neq 1$, the Prelec function (Definition 5) has exactly three fixed points, at 0, $p^* = e^{-\left(\frac{1}{\beta}\right)^{\frac{1}{\alpha-1}}}$ and 1. In particular, for $\beta = 1$, $p^* = e^{-1}$.*

Proof of Lemma 2: From Propositions 1a and 3 it follows that 0 and 1 are fixed points of the Prelec function. For $\alpha \neq 1$ and $p^* \in (0, 1)$, a simple calculation shows that $e^{-\beta(-\ln p^*)^\alpha} = p^* \Leftrightarrow p^* = e^{-\left(\frac{1}{\beta}\right)^{\frac{1}{\alpha-1}}}$. In particular, $\beta = 1$ gives $p^* = e^{-1}$. ■.

Lemma 3 : Let $w(p)$ be the Prelec function (Definition 5) and let

$$f(p) = \alpha\beta(-\ln p)^\alpha + \ln p + 1 - \alpha, p \in (0, 1), \quad (12.1)$$

then

$$f'(p) = \frac{1}{p} [1 - \alpha^2\beta(-\ln p)^{\alpha-1}], p \in (0, 1), \quad (12.2)$$

$$w''(p) = \frac{w'(p)}{p(-\ln p)} f(p), p \in (0, 1), \quad (12.3)$$

$$w''(p) \lesseqgtr 0 \Leftrightarrow f(p) \lesseqgtr 0, p \in (0, 1). \quad (12.4)$$

Proof of Lemma 3: Differentiate (12.1) to get (12.2). Differentiate (5.2) twice and use (12.1) to get (12.3). $-\ln p > 0$, since $p \in (0, 1)$. $w'(p) > 0$ follows from Definitions (1) and (5) and Proposition (3). (12.4) then follows from (12.3). ■.

Lemma 4 : Let $w(p)$ be the Prelec function (Definition 5). Suppose $\alpha \neq 1$. Then

- (a) $w''(\tilde{p}) = 0$ for some $\tilde{p} \in (0, 1)$ and, for any such \tilde{p} :
 - (i) for $\alpha < 1$, $p < \tilde{p} \Rightarrow w''(p) < 0$ and $p > \tilde{p} \Rightarrow w''(p) > 0$.
 - (ii) for $\alpha > 1$, $p < \tilde{p} \Rightarrow w''(p) > 0$ and $p > \tilde{p} \Rightarrow w''(p) < 0$.
- (b) The Prelec function has a unique inflexion point, $\tilde{p} \in (0, 1)$, and is characterized by $f(\tilde{p}) = 0$, i.e., $\alpha\beta(-\ln \tilde{p})^\alpha + \ln \tilde{p} + 1 - \alpha = 0$.
- (c) $\beta = 1 \Rightarrow \tilde{p} = e^{-1}$.
- (d) $\frac{\partial \tilde{p}}{\partial \beta} = \frac{\alpha \tilde{p} (-\ln \tilde{p})^{1+\alpha}}{(\alpha-1)(\alpha-\ln \tilde{p})}$.
- (e) $\frac{\partial [w(\tilde{p})-\tilde{p}]}{\partial \beta} = \frac{\tilde{p}(-\ln \tilde{p})^{1+\alpha}}{(\alpha-1)(\alpha-\ln \tilde{p})} \left(e^{\frac{1-\alpha}{\alpha}} (\tilde{p})^{\frac{1-\alpha}{\alpha}} - \alpha \right)$.
- (f) $\tilde{p} \lesseqgtr w(\tilde{p}) \Leftrightarrow \beta \lesseqgtr 1$.

Proof of Lemma 4: (a) Suppose $\alpha < 1$. From (12.1) we see that $f(p) \rightarrow 1 - \alpha > 0$, as $p \rightarrow 1$. We also see that $f(p) = \left[\frac{\alpha\beta}{(-\ln p)^{1-\alpha}} + 1 \right] \ln p + 1 - \alpha \rightarrow -\infty$ as $p \rightarrow 0$. Since f is continuous, it follows that $f(\tilde{p}) = 0$, for some $\tilde{p} \in (0, 1)$. From (12.4), it follows that $w''(\tilde{p}) = 0$. Since $\alpha < 1$, (12.1) gives $\alpha\beta(-\ln \tilde{p})^\alpha + \ln \tilde{p} < 0$ and, hence,

$$\tilde{p} < e^{-(\alpha\beta)^{\frac{1}{1-\alpha}}}. \quad (12.5)$$

Consider the case, $p < \tilde{p}$. From (12.5) it follows that $p < e^{-(\alpha\beta)^{\frac{1}{1-\alpha}}}$ and, hence, $1 - \frac{\alpha^2\beta}{(-\ln p)^{1-\alpha}} > 1 - \alpha > 0$. Hence, from (12.2), $f'(p) > 0$. Since $f(\tilde{p}) = 0$, it follows that $f(p) < 0$. Hence, from (12.4), it follows that $w''(p) < 0$. This establishes $p < \tilde{p} \Rightarrow$

$w''(p) < 0$. The derivation of the second part of Lemma 4(ai) is similar. The case $\alpha > 1$ is similar.

(b) follows from (a) and (12.1), (12.4).

(c) Since $f(e^{-1}) = 0$ for $\beta = 1$, it follows from (b) that $\tilde{p} = e^{-1}$ in this case.

(d) Differentiating the identity $f(\tilde{p}) = 0$ with respect to β gives $\frac{\partial \tilde{p}}{\partial \beta} = \frac{\alpha \tilde{p}(-\ln \tilde{p})^\alpha}{\alpha^2 \beta (-\ln \tilde{p})^{\alpha-1} - 1}$,

then using $f(\tilde{p}) = 0$ gives $\frac{\partial \tilde{p}}{\partial \beta} = \frac{\alpha \tilde{p}(-\ln \tilde{p})^{1+\alpha}}{(\alpha-1)(\alpha-\ln \tilde{p})}$.

(e) Differentiate $w(\tilde{p}) - \tilde{p} = e^{-\beta(-\ln \tilde{p})^\alpha} - \tilde{p}$ with respect to β , and use (d) and $f(\tilde{p}) = 0$, to get $\frac{\partial[w(\tilde{p})-\tilde{p}]}{\partial \beta} = \frac{\tilde{p}(-\ln \tilde{p})^{1+\alpha}}{(\alpha-1)(\alpha-\ln \tilde{p})} \left(e^{\frac{1-\alpha}{\alpha}} (\tilde{p})^{\frac{1-\alpha}{\alpha}} - \alpha \right)$.

(f) Assume $\alpha < 1$. For $\beta = 1$, $\tilde{p} = e^{-1}$ and, hence, $e^{\frac{1-\alpha}{\alpha}} (\tilde{p})^{\frac{1-\alpha}{\alpha}} - \alpha = e^{\frac{1-\alpha}{\alpha}} (e^{-1})^{\frac{1-\alpha}{\alpha}} - \alpha = 1 - \alpha > 0$. Since $\frac{\partial \tilde{p}}{\partial \beta} < 0$ for $\alpha < 1$, it follows that $e^{\frac{1-\alpha}{\alpha}} (\tilde{p})^{\frac{1-\alpha}{\alpha}} - \alpha > 0$ for $\beta \leq 1$. Hence, $\frac{\partial[w(\tilde{p})-\tilde{p}]}{\partial \beta} < 0$ for $\beta \leq 1$. We have $w(\tilde{p}) - \tilde{p} = w(e^{-1}) - e^{-1} = w(p^*) - p^* = 0$ for $\beta = 1$ (recall part c and Lemma 2). Hence, $w(\tilde{p}) > \tilde{p}$ for $\beta < 1$. The case $\beta \geq 0$ is similar. The case $\alpha > 1$ is similar. ■.

Proposition 13 : Suppose $\alpha \neq 1$. Then:

(a) The Prelec function (Definition 5) has exactly three fixed points at, respectively, 0, $p^* = e^{-\left(\frac{1}{\beta}\right)^{\frac{1}{\alpha-1}}}$ and 1.

(b) The Prelec function has a unique inflexion point, $\tilde{p} \in (0, 1)$, at which $w''(\tilde{p}) = 0$.

(c) If $\alpha < 1$, the Prelec function is strictly concave for $p < \tilde{p}$ and strictly convex for $p > \tilde{p}$ (inverse S-shaped).

(d) If $\alpha > 1$, the Prelec function is strictly convex for $p < \tilde{p}$ and strictly concave for $p > \tilde{p}$ (S-shaped).

(e) If $\beta < 1$, the inflexion point, \tilde{p} , lies above the 45° line ($\tilde{p} < w(\tilde{p})$).

(f) If $\beta = 1$, the inflexion point, \tilde{p} , lies on the 45° line ($\tilde{p} = w(\tilde{p})$).

(g) If $\beta > 1$, the inflexion point, \tilde{p} , lies below the 45° line ($\tilde{p} > w(\tilde{p})$).

Proof of Proposition 13: (a) is established by Lemma 2. (b) is established by Lemma 4b. (c) follows from Lemma 4a(i). (d) follows from Lemma 4a(ii). (e), (f) and (g) follow from Lemma 4f ■.

Proof of Corollary 1: Immediate from Proposition 13. ■.

Table 1, below, exhibits the various cases established by Proposition 13.

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

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.

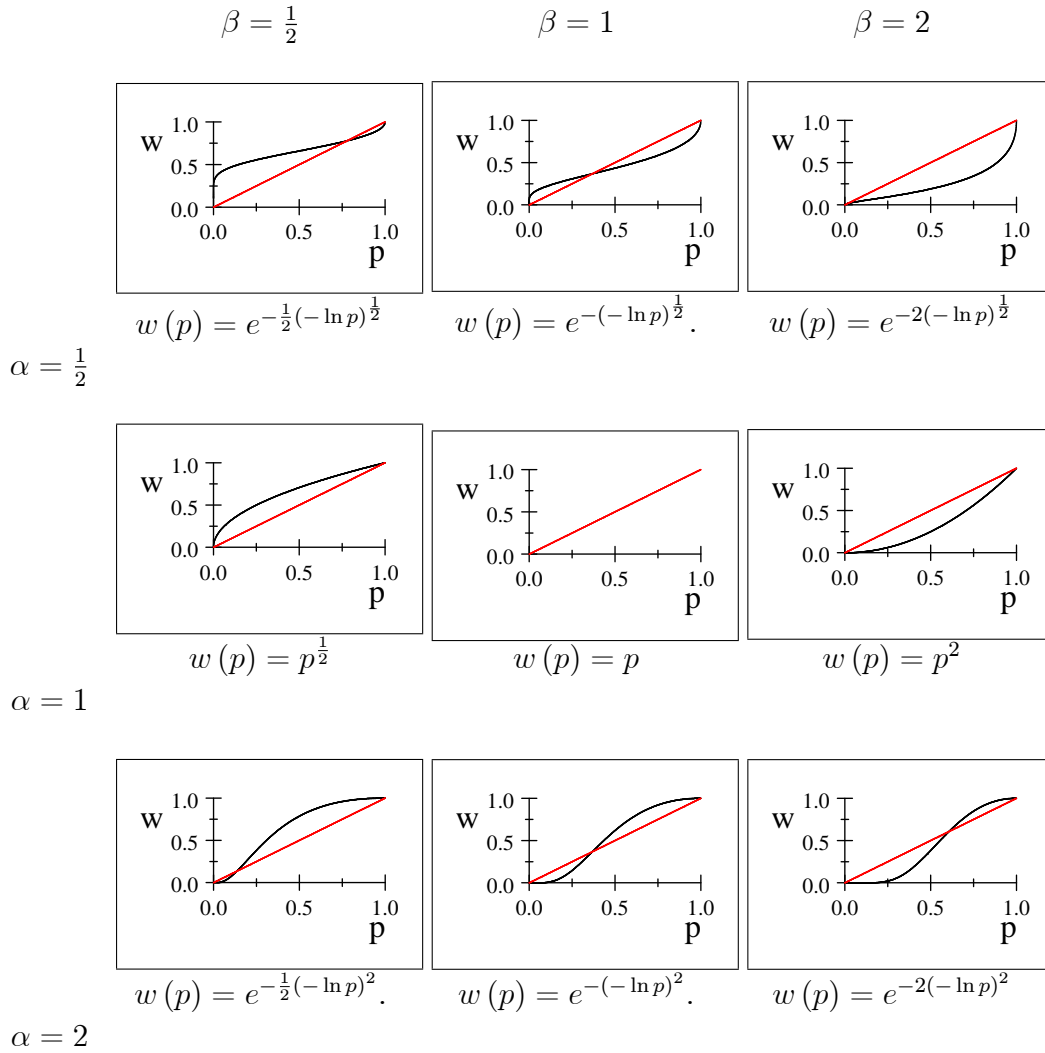


Table 2: Representative graphs of $w(p) = e^{-\beta(-\ln p)^\alpha}$.

Corollary 1 : Suppose $\alpha \neq 1$. Then $\tilde{p} = p^* = e^{-1}$ (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 5.1).
- (b) If $\alpha > 1$, then w is strictly convex for $p < e^{-1}$ and strictly concave for $p > e^{-1}$ (S shape, see Figure 5.2).

12.2. Cauchy's algebraic functional equations.

We start with Cauchy's first algebraic functional equation, with its classic proof. Our notation is standard. In particular: \mathbb{R} : reals, \mathbb{R}_+ : non-negative reals, \mathbb{R}_{++} : positive reals

and C^1 : class of continuous functions with continuous first derivatives.

12.2.1. Cauchy's first algebraic functional equation

Theorem 1 : If $f : \mathbb{R} \rightarrow \mathbb{R}$ is continuous and satisfies $\forall x \in \mathbb{R}, \forall y \in \mathbb{R}, f(x+y) = f(x) + f(y)$, then $\exists c \in \mathbb{R}, \forall x \in \mathbb{R}, f(x) = cx$.

Proof. By mathematical induction it follows that, $\forall n \in \mathbb{N}, \forall x_1, x_2, \dots, x_n \in \mathbb{R}, f(\sum_{i=1}^n x_i) = \sum_{i=1}^n f(x_i)$. In particular, $\forall n \in \mathbb{N}, \forall x \in \mathbb{R}, f(nx) = nf(x)$. Let $n \in \mathbb{N}, x \in \mathbb{R}$. Let $y = \frac{1}{n}x$. Then $x = ny$. Hence, $f(x) = f(ny) = nf(y) = nf(\frac{1}{n}x)$. Thus, $\forall n \in \mathbb{N}, \forall x \in \mathbb{R}, f(\frac{1}{n}x) = \frac{1}{n}f(x)$. And, so, $\forall m, n \in \mathbb{N}, \forall y \in \mathbb{R}, f(\frac{m}{n}y) = \frac{1}{n}f(my) = \frac{m}{n}f(y)$. From the continuity of f it follows that $\forall x, y \in \mathbb{R}, f(xy) = xf(y)$. In particular, for $y = 1$, we get $\forall x \in \mathbb{R}, f(x) = xf(1)$. Letting $c = f(1)$, we get $\forall x \in \mathbb{R}, f(x) = cx$. ■

Remark 7 : Note that the rational number, $\frac{m}{n}$, can be arbitrarily large. Hence, for the proof to go through, we do need $f : \mathbb{R} \rightarrow \mathbb{R}$ (or $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ or $f : \mathbb{R}_{++} \rightarrow \mathbb{R}_{++}$). In particular, this proof is not valid for the case $f : (a, b) \rightarrow \mathbb{R}$ when (a, b) is a bounded interval; which is what we need. This is why we need a local form of this theorem.

12.2.2. Local forms of Cauchy's algebraic functional equations.

The third Cauchy equation arises naturally in the course of our proof of Proposition 14. However, we need a local form of it. That is, a form restricted to a bounded real interval around zero, rather than the whole real line (recall Remark7). We achieve this by replacing the assumption of continuity with the stronger assumption of differentiability.³⁴ As in the classical approach, we give the proof for the first equation. We then transform the third equation to the second which, in turn, we transform to the first. There is a fourth Cauchy algebraic functional equation, but we need not be concerned with it here.

Theorem 2 : Let $f : (a, b) \rightarrow \mathbb{R}$ be C^1 , $a < 0 < b$, $\forall x, y \in (a, b)$, s.t. $x + y \in (a, b)$, $f(x+y) = f(x) + f(y)$. Then $\exists c \in \mathbb{R}, \forall x \in (a, b), f(x) = cx$.

Proof. $f(0) = f(0+0) = f(0) + f(0)$. Hence, $f(0) = 0$.
 $\frac{f(x+\delta x)-f(x)}{\delta x} = \frac{f(x)+f(\delta x)-f(x)}{\delta x} = \frac{f(\delta x)}{\delta x}$. Hence, $f'(x) = \lim_{\delta x \rightarrow 0} \frac{f(x+\delta x)-f(x)}{\delta x} = \lim_{\delta x \rightarrow 0} \frac{f(\delta x)}{\delta x}$, which is independent of x , thus, $\exists c \in \mathbb{R}, f'(x) = c$ and so $f(x) = cx + C$. Hence, $0 = f(0) = C$. It follows that $\exists c \in \mathbb{R}, f(x) = cx$. ■

Theorem 3 : Let $g : (A, B) \rightarrow \mathbb{R}$ be C^1 , $0 < A < 1 < B$, $\forall X, Y \in (A, B)$, s.t. $XY \in (A, B)$, $g(XY) = g(X) + g(Y)$. Then $\exists c \in \mathbb{R}, \forall X \in (A, B), g(X) = c \ln X$.

³⁴Theorem 2 (and hence also Theorems 3 and 4) holds if f is either continuous or monotonic (see, for example, Theorem 2.6.9 in Eichhorn 1987). The assumption of differentiability, however, facilitates a much shorter proof.

Proof. Let $a_i = \ln A$, $b_i = \ln B$. Then $a_i < 0 < b_i$, $x \in (a_i, b_i) \Leftrightarrow e^x \in (A, B)$ and $x, y, x + y \in (a_i, b_i) \Rightarrow e^x, e^y, e^x e^y = e^{x+y} \in (A, B)$.

Define $f : (a_i, b_i) \rightarrow \mathbb{R}$ by $\forall x \in (a_i, b_i)$, $f(x) = g(e^x)$. Since $e^x \in (A, B)$, f is well defined. Since g is C^1 , f is also C^1 and for $x, y, x + y \in (a_i, b_i)$, $f(x + y) = g(e^{x+y}) = g(e^x e^y) = g(e^x) + g(e^y) = f(x) + f(y)$. Hence, from Theorem 1, $\exists c \in \mathbb{R}$, $\forall x \in (a_i, b_i)$, $f(x) = cx$. Hence, $\forall x \in (a_i, b_i)$, $g(e^x) = cx$. Let $X \in (A, B)$, $x = \ln X$, then $x \in (a_i, b_i)$, $g(X) = g(e^x) = cx = c \ln X$, i.e., $\forall X \in (A, B)$, $g(X) = c \ln X$. ■

Theorem 4 : Let $G : (A, B) \rightarrow \mathbb{R}_{++}$ is C^1 , $0 < A < 1 < B$, $\forall X, Y \in (A, B)$, s.t. $XY \in (A, B)$, $G(XY) = G(X)G(Y)$. Then $\exists c \in \mathbb{R}$, $\forall X \in (A, B)$, $G(X) = X^c$.

Proof. Define $g : (A, B) \rightarrow \mathbb{R}$ by $\forall X \in (A, B)$, $g(X) = \ln G(X)$. Since $G(X) > 0$, g is well defined. Since G is C^1 , g is also C^1 and for $X, Y \in (A, B)$, s.t. $XY \in (A, B)$, $g(XY) = \ln G(XY) = \ln(G(X)G(Y)) = \ln G(X) + \ln G(Y) = g(X) + g(Y)$. Hence, from Theorem 3, $\exists c \in \mathbb{R}$, $\forall X \in (A, B)$, $g(X) = c \ln X$. Hence, $\forall X \in (A, B)$, $\ln G(X) = c \ln X = \ln X^c$. Hence, $\forall X \in (A, B)$, $G(X) = X^c$. ■

12.3. Axiomatic foundations of the composite Prelec probability weighting function.

Prelec (1998) gave an axiomatic derivation of (5.1) and (5.2) based on ‘compound invariance’, Luce (2001) provided a derivation based on ‘reduction invariance’ and al-Nowaihi and Dharni (2006) gave a derivation based on ‘power invariance’. Since the Prelec function satisfies all three, ‘compound invariance’, ‘reduction invariance’ and ‘power invariance’ are all equivalent.

Here we introduced a version of *power invariance* that we call *local power invariance*. On the basis of this behavioral property, we shall axiomatically derive the *composite Prelec function* (CPF); see Proposition 14, below.

Definition 9 (*Composite Prelec function, CPF*): By the *composite Prelec function* we mean the function $w : [0, 1] \rightarrow [0, 1]$ given by

$$w(p) = \begin{cases} 0 & \text{if } p = 0 \\ e^{-\beta_i(-\ln p)^{\alpha_i}} & \text{if } p_{i-1} < p \leq p_i, i = 1, 2, \dots, n, \end{cases} \quad (12.6)$$

where $\alpha_i > 0$, $\beta_i > 0$, $p_0 = 0$, $p_n = 1$ and

$$e^{-\beta_i(-\ln p_i)^{\alpha_i}} = e^{-\beta_{i+1}(-\ln p_i)^{\alpha_{i+1}}}, i = 1, 2, \dots, n - 1. \quad (12.7)$$

The restriction (12.7) is needed to insure that w is continuous.

Definition 10 (Power invariance; al-Nowaihi and Dhimi, 2006): A probability weighting function, w , satisfies power invariance if: $\forall p, q \in (0, 1)$, $(w(p))^\mu = w(q) \Rightarrow (w(p^\lambda))^\mu = w(q^\lambda)$, $\lambda, \mu \in \{2, 3\}$.

Definition 11 (Local power invariance): Let $0 = p_0 < p_1 < \dots < p_n = 1$. A probability weighting function, $w(\cdot)$, satisfies local power invariance if, for $i = 1, 2, \dots, n$, $w(\cdot)$ is C^1 on (p_{i-1}, p_i) and $\forall p, q \in (p_{i-1}, p_i)$, $(w_i(p))^\mu = w_i(q)$ and $p^\lambda, q^\lambda \in (p_{i-1}, p_i)$ imply $(w(p^\lambda))^\mu = w(q^\lambda)$.

Definition 12 (Notation): Let $0 = p_0 < p_1 < \dots < p_n = 1$. Define $P_1 = (0, p_1]$, $P_n = [p_{n-1}, 1)$ and $P_i = [p_{i-1}, p_i]$, $i = 2, 3, \dots, n-1$. Given $p \in P_i$, $i = 1, 2, \dots, n$, define Λ_i as follows. $\Lambda_1 = [\frac{\ln p_1}{\ln p}, \infty)$, $\Lambda_n = (0, \frac{\ln p_{n-1}}{\ln p}]$, $\Lambda_i = [\frac{\ln p_i}{\ln p}, \frac{\ln p_{i-1}}{\ln p}]$, $i = 2, 3, \dots, n-1$.

Lemma 5 : Let p_i , P_i and Λ_i be as in Definition 12. Then,

$$\text{Let } p \in (p_{i-1}, p_i) . \text{ Then } p^\lambda \in (p_{i-1}, p_i) \Leftrightarrow \lambda \in \left(\frac{\ln p_i}{\ln p}, \frac{\ln p_{i-1}}{\ln p} \right), \quad (12.8)$$

$$\text{furthermore, } 0 < \frac{\ln p_i}{\ln p} < 1 < \frac{\ln p_{i-1}}{\ln p}. \quad (12.9)$$

$$\text{Let } p \in P_i . \text{ Then } p^\lambda \in P_i \Leftrightarrow \lambda \in \Lambda_i. \quad (12.10)$$

Proposition 14 (CPF representation): The following are equivalent.

- (a) The probability weighting function, w , satisfies local power invariance.
- (b) There are functions, $\varphi_i : \Lambda_i \rightarrow \mathbb{R}_{++}$, such that φ_i is C^1 on $\left(\frac{\ln p_i}{\ln p}, \frac{\ln p_{i-1}}{\ln p} \right)$, $i = 1, 2, \dots, n$, where $0 = p_0 < p_1 < \dots < p_n = 1$, and, $\forall p \in P_i$, $\forall \lambda \in \Lambda_i$, $w(p^\lambda) = (w(p))^{\varphi_i(\lambda)}$. Moreover, for each $i = 1, 2, \dots, n$, $\exists \alpha_i \in (0, \infty)$, $\varphi_i(\lambda) = \lambda^{\alpha_i}$.
- (c) w is a composite Prelec function (Definition 9).

Proof of Proposition 14 (CPF representation): (a) \Rightarrow (b). Suppose the probability weighting function, w , satisfies local power invariance.

Let

$$f(x, \lambda) = w\left(\left(w^{-1}(e^{-x})\right)^\lambda\right), \quad x, \lambda \in \mathbb{R}_{++}, \quad (12.11)$$

and

$$\varphi(\lambda) = -\ln f(1, \lambda) = -\ln w\left(\left(w^{-1}(e^{-1})\right)^\lambda\right), \quad \lambda \in \mathbb{R}_{++}. \quad (12.12)$$

Clearly,

$$\varphi \text{ maps } \mathbb{R}_{++} \text{ into } \mathbb{R}_{++}. \quad (12.13)$$

Since $w^{-1}(e^{-1}) \in (0, 1)$, it follows that $(w^{-1}(e^{-1}))^\lambda$ is a strictly decreasing function of λ , and so are $w\left(\left(w^{-1}(e^{-1})\right)^\lambda\right)$ and $\ln w\left(\left(w^{-1}(e^{-1})\right)^\lambda\right)$. Hence, from (12.12),

$$\varphi \text{ is a strictly increasing function of } \lambda. \quad (12.14)$$

From (12.11) we get

$$f(-\mu \ln w(p), \lambda) = w\left(\left(w^{-1}\left((w(p))^\mu\right)\right)^\lambda\right), p \in (0, 1), \lambda, \mu \in \mathbb{R}_{++}. \quad (12.15)$$

Let $0 = p_0 < p_1 < \dots < p_n = 1$.

Since w is C^1 on (p_{i-1}, p_i) it follows, from (12.12) and (12.8), that

$$\varphi \text{ is } C^1 \text{ on } \left(\frac{\ln p_i}{\ln p}, \frac{\ln p_{i-1}}{\ln p}\right). \quad (12.16)$$

Let

$$p, q \in (p_{i-1}, p_i), (w(p))^\mu = w(q), p^\lambda, q^\lambda \in (p_{i-1}, p_i). \quad (12.17)$$

From (12.17) we get

$$q = w^{-1}\left((w(p))^\mu\right). \quad (12.18)$$

From (12.17) and local power invariance, we get

$$(w(p^\lambda))^\mu = w(q^\lambda). \quad (12.19)$$

Substituting for q from (12.18) into (12.19), we get

$$(w(p^\lambda))^\mu = w\left(\left(w^{-1}\left((w(p))^\mu\right)\right)^\lambda\right), p, p^\lambda \in (p_{i-1}, p_i). \quad (12.20)$$

From (12.20) and (12.15) we get

$$f(-\mu \ln w(p), \lambda) = (w(p^\lambda))^\mu, p, p^\lambda \in (p_{i-1}, p_i). \quad (12.21)$$

In particular, for $\mu = 1$, (12.21) gives

$$f(-\ln w(p), \lambda) = w(p^\lambda), p, p^\lambda \in (p_{i-1}, p_i). \quad (12.22)$$

From (12.22) we get

$$(f(-\ln w(p), \lambda))^\mu = (w(p^\lambda))^\mu, p, p^\lambda \in (p_{i-1}, p_i). \quad (12.23)$$

From (12.21) and (12.23) we get

$$f(-\mu \ln w(p), \lambda) = (f(-\ln w(p), \lambda))^\mu, p, p^\lambda \in (p_{i-1}, p_i). \quad (12.24)$$

Put

$$z = -\ln w(p). \quad (12.25)$$

From (12.24) and (12.25) we get

$$f(\mu z) = (f(z, \lambda))^\mu, p, p^\lambda \in (p_{i-1}, p_i). \quad (12.26)$$

From (12.12) and (12.26) we get

$$f(\mu) = (f(1, \lambda))^\mu = e^{-\mu\varphi(\lambda)}, p, p^\lambda \in (p_{i-1}, p_i), \quad (12.27)$$

and, hence,

$$f(-\ln w(p), \lambda) = (w(p))^{\varphi(\lambda)}, p, p^\lambda \in (p_{i-1}, p_i). \quad (12.28)$$

From (12.22) and (12.28) we get

$$w(p^\lambda) = (w(p))^{\varphi(\lambda)}, p, p^\lambda \in (p_{i-1}, p_i), \quad (12.29)$$

from which we get,

$$\varphi(\lambda) = \frac{\ln w(p^\lambda)}{\ln w(p)}, p, p^\lambda \in (p_{i-1}, p_i). \quad (12.30)$$

Let $p, p^\lambda, p^\mu, p^{\lambda\mu} \in (p_{i-1}, p_i)$. From (12.29) and (12.30) we get

$$\begin{aligned} \varphi(\lambda\mu) &= \frac{\ln w(p^{\lambda\mu})}{\ln w(p)} = \frac{\ln w((p^\mu)^\lambda)}{\ln w(p)} = \frac{\ln[(w(p^\mu))^{\varphi(\lambda)}]}{\ln w(p)} = \frac{\varphi(\lambda) \ln(w(p^\mu))}{\ln w(p)} = \frac{\varphi(\lambda) \ln[(w(p))^{\varphi(\mu)}]}{\ln w(p)} = \\ &= \frac{\varphi(\lambda)\varphi(\mu) \ln w(p)}{\ln w(p)} = \varphi(\lambda)\varphi(\mu), \text{ i.e.,} \end{aligned}$$

$$\varphi(\lambda\mu) = \varphi(\lambda)\varphi(\mu), p, p^\lambda, p^\mu, p^{\lambda\mu} \in (p_{i-1}, p_i). \quad (12.31)$$

From (12.8), (12.9), (12.16) and (12.31) we have: φ is C^1 on $\left(\frac{\ln p_i}{\ln p}, \frac{\ln p_{i-1}}{\ln p}\right)$, $0 < \frac{\ln p_i}{\ln p} < 1 < \frac{\ln p_{i-1}}{\ln p}$, $\forall \lambda, \mu \in \left(\frac{\ln p_i}{\ln p}, \frac{\ln p_{i-1}}{\ln p}\right)$, s.t. $\lambda\mu \in \left(\frac{\ln p_i}{\ln p}, \frac{\ln p_{i-1}}{\ln p}\right)$, $\varphi(\lambda\mu) = \varphi(\lambda)\varphi(\mu)$. Hence, by Theorem 4 (see Appendix 2, below), we have,

$$\exists \alpha_i \in \mathbb{R}, \forall \lambda \in \left(\frac{\ln p_i}{\ln p}, \frac{\ln p_{i-1}}{\ln p}\right), \varphi(\lambda) = \lambda^{\alpha_i}. \quad (12.32)$$

But, by (12.14), φ is a strictly increasing function of λ . Hence,

$$\alpha_i > 0. \quad (12.33)$$

Let P_i and Λ_i be as in Definition 12. Let $p \in P_i$. Define $\varphi_i : \Lambda_i \rightarrow \mathbb{R}_{++}$ by $\varphi_i(\lambda) = \lambda^{\alpha_i}$. Then, clearly, φ_i is C^1 on $\left(\frac{\ln p_i}{\ln p}, \frac{\ln p_{i-1}}{\ln p}\right)$. A simple calculation verifies that $\forall p \in P_i, \forall \lambda \in \Lambda_i$, $w(p^\lambda) = (w(p))^{\varphi_i(\lambda)}$. This completes the proof of part (b).

(b) \Rightarrow (c). Since $e^{-1} \in (0, 1)$, $e^{-1} \in P_i$ for some $i = 1, 2, \dots, n$. We first establish the result for P_i , then we use induction, and the continuity conditions (12.7), to extend the result to $P_{i+1}, P_{i+2}, \dots, P_n$ and $P_{i-1}, P_{i-2}, \dots, P_1$. Let $\beta_i = -\ln w(e^{-1})$. Then $w(e^{-1}) = e^{-\beta_i}$. Let $p \in P_i$. Let $\lambda = -\ln p$. Then $p = e^{-\lambda}$. Hence $w(p) = w(e^{-\lambda}) = w((e^{-1})^\lambda) = (w(e^{-1}))^{\varphi_i(\lambda)} = (e^{-\beta_i})^{\lambda^{\alpha_i}} = e^{-\beta_i \lambda^{\alpha_i}} = e^{-\beta_i (-\ln p)^{\alpha_i}}$. Thus we have shown

$$w(p) = e^{-\beta_i (-\ln p)^{\alpha_i}}, p \in P_i. \quad (12.34)$$

Let $p \in P_i$. Let $\lambda = \frac{\ln p}{\ln p_i}$. Then $p = p_i^\lambda$. Hence, $w(p) = w(p_i^\lambda) = (w(p_i))^{\varphi_{i+1}(\lambda)} = (w(p_i))^{\lambda^{\alpha_{i+1}}} = (e^{-\beta_i(-\ln p_i)^{\alpha_i}})^{\lambda^{\alpha_{i+1}}} = (e^{-\beta_{i+1}(-\ln p_i)^{\alpha_{i+1}}})^{\lambda^{\alpha_{i+1}}} = e^{-\beta_{i+1}(-\lambda \ln p_i)^{\alpha_{i+1}}} = e^{-\beta_{i+1}(-\ln p_i^\lambda)^{\alpha_{i+1}}} = e^{-\beta_{i+1}(-\ln p)^{\alpha_{i+1}}}$. Thus we have shown

$$w(p) = e^{-\beta_{i+1}(-\ln p)^{\alpha_{i+1}}}, p \in P_{i+1}. \quad (12.35)$$

Let $p \in P_{i-1}$. Let $\lambda = \frac{\ln p}{\ln p_{i-1}}$. Then $p = p_{i-1}^\lambda$. Hence, $w(p) = w(p_{i-1}^\lambda) = (w(p_{i-1}))^{\varphi_{i-1}(\lambda)} = (w(p_{i-1}))^{\lambda^{\alpha_{i-1}}} = (e^{-\beta_i(-\ln p_{i-1})^{\alpha_i}})^{\lambda^{\alpha_{i-1}}} = (e^{-\beta_{i-1}(-\ln p_{i-1})^{\alpha_{i-1}}})^{\lambda^{\alpha_{i-1}}} = e^{-\beta_{i-1}(-\lambda \ln p_{i-1})^{\alpha_{i-1}}} = e^{-\beta_{i-1}(-\ln p_i^\lambda)^{\alpha_{i-1}}} = e^{-\beta_{i-1}(-\ln p)^{\alpha_{i-1}}}$. Thus we have shown

$$w(p) = e^{-\beta_{i-1}(-\ln p)^{\alpha_{i-1}}}, p \in P_{i-1}. \quad (12.36)$$

Continuing the above process, we get

$$w(p) = e^{-\beta_i(-\ln p)^{\alpha_i}}, p \in P_i, i = 1, 2, \dots, n, \quad (12.37)$$

which establishes part (c).

Finally, a simple calculation shows that (c) implies (a). ■

References

- [1] al-Nowaihi, A., Dhimi, S., (2010) Composite Prospect Theory: A proposal to combine ‘prospect theory’ and ‘cumulative prospect theory.’ University of Leicester Discussion Paper.
- [2] al-Nowaihi, A., Dhimi, S., 2006. A simple derivation of Prelec’s probability weighting function. *The Journal of Mathematical Psychology* 50 (6), 521-524.
- [3] al-Nowaihi, A., Bradley, I. and Dhimi, S. (2008) ‘The Utility Function Under Prospect Theory’, *Economics Letters* 99, p.337–339.
- [4] Bar Ilan, Avner and Bruce Sacerdote (2004) “The response of criminals and noncriminals to fines” *Journal of Law and Economics*, Vol. 47, 1-17.
- [5] Camerer, C. and Kunreuther, H. (1989). Experimental markets for insurance. *Journal of Risk and Uncertainty*, 2, 265-300.
- [6] Dhimi, S. and al-Nowaihi, A. (2010a) The behavioral economics of crime and punishment. University of Leicester Discussion Paper.
- [7] Dhimi, S., al-Nowaihi, A., 2007. Why do people pay taxes? Prospect theory versus expected utility theory. *Journal of Economic Behavior and Organization*, 64, 171–192.

- [8] Dhami, S., al-Nowaihi, A., 2010b. Optimal income taxation in the presence of tax evasion: Expected utility versus prospect theory. Forthcoming in *Journal of Economic Behavior and Organization*.
- [9] Eichhorn W. (1978). *Functional Equations in Economics*. Addison-Wesley.
- [10] Frank, C.R., 1965. Entry in a Cournot Market. *Review of Economic Studies* 32, 245-250.
- [11] Gonzalez, R., Wu, G., 1999. On the shape of the probability weighting function. *Cognitive Psychology* 38, 129-166.
- [12] Kahneman, D., and Tversky, A. (Eds.). (2000). *Choices, values and frames*. New York: Cambridge University Press.
- [13] Kahneman D., Tversky A., 1979. Prospect theory : An analysis of decision under risk. *Econometrica* 47, 263-291.
- [14] Kunreuther, H., Ginsberg, R., Miller, L., Sagi, P., Slovic, P., Borkan, B., Katz, N., 1978. *Disaster insurance protection: Public policy lessons*, Wiley, New York.
- [15] 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.
- [16] Luce, R. D., 2001. Reduction invariance and Prelec's weighting functions. *Journal of Mathematical Psychology* 45, 167-179.
- [17] 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).
- [18] McClelland, G. H., W. D. Schulze, and D. L. Coursey (1993): Insurance for Low-Probability Hazards; a Bimodal Response to Unlikely Events, in: *Journal of Risk and Uncertainty*, 7: 95-116.
- [19] Novshek, W., 1980. Cournot equilibrium with free entry. *Review of Economic Studies* XLVII, 473-486.
- [20] Plunkett 2010. *Plunkett's Insurance Industry Almanac*, Plunkett Research, Ltd.
- [21] 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.

- [22] Prelec, D., 1998. The probability weighting function. *Econometrica* 60, 497-528.
- [23] Quiggin, J., 1982. A theory of anticipated utility. *Journal of Economic Behavior and Organization* 3, 323-343.
- [24] Quiggin, J., 1993. *Generalized Expected Utility Theory*, Kluwer Academic Publishers.
- [25] Rieger, M.O., Wang, M., 2006. Cumulative prospect theory and the St. Petersburg paradox. *Economic Theory* 28, 665-679.
- [26] Royal Society for the prevention of accidents (2005) The risk of using a mobile phone while driving. Department for Transport, Local Government and the Regions.
- [27] Ruffin, R.J., 1971. Cournot oligopoly and competitive behavior. *Review of Economic Studies* 38, 493-502.
- [28] Schade, C., Kunreuther, H., and Kaas, K. P. (2001) Low-Probability Insurance: Are Decisions Consistent with Normative Predictions? mimeo. University of Humboldt.
- [29] Chris Starmer, 2000. "Developments in Non-expected Utility Theory: The Hunt for a Descriptive Theory of Choice under Risk," *Journal of Economic Literature*, 38: 332-382.
- [30] Tversky, A., Kahneman D., 1992. Advances in prospect theory : Cumulative representation of uncertainty. *Journal of Risk and Uncertainty* 5, 297-323.
- [31] 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.