The behavioral economics of the demand for insurance

Ali al-Nowaihi† Sanjit Dhami‡

20 July 2010

Abstract

We focus on the following findings from insurance demand. When faced with low probability losses, (1) some decision makers buy inadequate insurance, while (2) others are eager to buy insurance even when such insurance is not good value. Furthermore, (3) keeping the expected loss fixed, there is a probability below which the take-up of insurance drops dramatically, and (4) the proportion of decision makers that are split between (1) and (2) depends on the context and the frame of the problem. Expected utility (EU) fails on 1-4. It’s main alternatives, rank dependent utility (RDU) and cumulative prospect theory (CP) satisfy 2 but fail on 1, 3, 4. We use a new class of axiomatically-founded probability weighting functions, the composite Prelec weighting functions (CPF). When the CPF is combined with CP, we get composite prospect theory (CCP). CCP is able to simultaneously account for 1-4.

Keywords: Decision making under risk, Insurance, Context dependence and framing, Composite Prelec probability weighting functions, Composite cumulative prospect theory.

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

---

*We are very grateful for insightful comments made at the presentation of this paper at the 2010 FUR XIV conference in Newcastle and seminars at the University of Leicester and ISET in Georgia. Dr. Sanjit Dhami would also like to acknowledge a one semester study leave from the University of Leicester.

†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.
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. Theoretical models have mainly used the expected utility framework (EU) to model insurance demand. Yet, despite impressive progress, we show that existing theoretical models are unable to explain the stylized facts on the take-up of insurance, particularly when losses arise with low probability. The main motivation for the paper is to address these concerns.

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

There is overwhelming evidence against expected utility theory (EU). It is, thus, not surprising that the application of EU to insurance is problematic. 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. However, EU is unable to explain several stylized facts from insurance, as we now outline.

First, Kunreuther et al. (1978) gives experimental and field evidence that is inconsistent with the predicted take-up of insurance against low probability losses under EU (more on this below). 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 very low probability events; see Kunreuther et al (1978). 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 to either. This is contradicted by the experimental evidence (Kahneman and Tversky, 1979: 269-271).

1This evidence is well documented. It includes, failure of the independence axiom, implausible attitudes to risk for small and large stakes (Rabin’s calibration theorem), preference reversals, failure to incorporate the psychologically important phenomena of loss aversion, reference dependence and non-linear probability weighting etc; see Kahneman and Tversky (1979), Kahneman and Tversky (2000), and Starmer (2000).

2For 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.
1.2. A brief note on non-expected utility (non-EU) theories

The problems with EU have motivated several non-expected utility (non-EU) alternatives that are more successful. The most important are Quiggin’s (1982, 1993) rank dependent utility theory (RDU); Kahneman and Tversky’s (1979) prospect theory (PT), and Tversky and Kahneman’s (1992) cumulative prospect theory (CP).3 CP can explain the anomalous result arising from probabilistic insurance that EU cannot.4

All non-EU theories postulate a (strictly increasing) probability weighting function. It is typically the case that \( w(p) : [0, 1] \rightarrow [0, 1] \), so \( w(0) = 0 \) and \( w(1) = 1 \); \( w(p) \) captures the subjective weight placed by decision makers on the objective probability, \( p \).

Example 1 (RDU): Consider the lottery \( L = (x_1, p_1; x_2, p_2) \) where outcome \( x_i \) occurs with probability \( p_i \geq 0 \), \( x_1 < x_2 \), \( p_1 + p_2 = 1 \).5 Let \( u(x) \) be the utility of \( x \), so under EU, \( EU(L) = p_1u(x_1) + p_2u(x_2) \). Under non-EU theories, decision makers use decision weights, \( \pi_i \), to evaluate the value of the lottery as \( V(L) = \pi_1u(x_1) + \pi_2u(x_2) \). In most modern non-EU theories, the decision weights are constructed as in rank dependent utility (RDU), by using cumulative transformations of probability.6 For the lottery, \( L \), under RDU,

\[
\pi_1 = w(p_1 + p_2) - w(p_2) = 1 - w(p_2), \quad \pi_2 = w(p_2) \quad [\vdots \ w(1) = 1]
\]

Example 2 (CP):7 Under cumulative prospect theory (CP), utility arises, not from the level of wealth (as under EU, RDU), but from changes relative to a reference point, \( x \).8 Let \( x_1 < x < x_2 \), then the decision maker is in the ‘domain of loss’ if outcome \( x_1 \) occurs and in the ‘domain of gains’ if outcome \( x_2 \) occurs. Thus, under CP, the lottery \( L = (x_1, p_1; x_2, p_2) \) is viewed as the ‘prospect’ \( \tilde{L} = (y_1, p_1; y_2, p_2) \) where \( y_1 = x_1 - x < 0 \) (losses) and \( y_2 = x_2 - x > 0 \) (gains). Under CP, the utility function \( v(y) \) satisfies the following: \( v(0) = 0 \) (reference dependence); for \( y > 0 \), \( -v(-y) > v(y) \) (loss aversion); and \( v(y) \) is strictly concave for gains and strictly convex for losses (which explains simultaneous insurance and gambling). An example of \( v(y) \) is given by (7.1), below. The decision weights under

---

3Mark 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 can explain everything that RDU can explain but the converse is false.

4In particular, these non-EU theories can explain the problems mentioned with EU theories in the opening footnote of subsection 1.1; see, for instance, Kahneman and Tversky (2000), and Starmer (2000).

5Everything we say here can be extended to general lotteries.

6This ensures that decision makers would not choose stochastically dominated options. This solved, at that time, an important problem, namely, that point transformations of probability can lead to stochastically dominated choices, as under Kahneman and Tversky’s (1979) prospect theory, PT. All subsequent mainstream non-EU theories in economics adopt this insight.


8The reference point, \( x \), is often (but not necessarily) taken to be the status quo.
CP are formed as under RDU in (1.1) but applied separately to the domain of gains and losses. The value of the prospect $\tilde{L}$ under CP is $V(\tilde{L}) = w(p_1)v(y_1) + w(p_2)v(y_2)$.

Based on a great deal of evidence, all mainstream non-EU theories in economics assume that the shape of $w$ is an inverse-S shape. We list this as stylized fact S1.

**S1: Decision makers overweight low probabilities but underweight high probabilities.**

A popular $w(p)$ function under non-EU theories, that is consistent with S1, is parsimonious and has axiomatic foundations, is the standard Prelec (1998) function: $w(p) = e^{-\beta(-\ln p)^{\alpha}}$, $0 < \alpha < 1$, $0 < \beta$.

Figure 1.1 plots it for the case $\alpha = 0.5$, $\beta = 1$. An important implication of S1 is given in Remark 1.

**Remark 1** For any of the standard probability weighting functions (e.g. the standard Prelec function) used in RDU and CP, decision makers infinitely overweight infinitesimal probabilities in the sense that $\lim_{p \to 0} w(p)/p = \infty$ and infinitely underweights near one probabilities in the sense $\lim_{p \to 1} [1 - w(p)]/[1 - p] = \infty$.

### 1.3. Insurance for low probability events-I

The seminal study of Kunreuther et al. (1978), with 135 expert contributors, provides striking evidence of individuals buying inadequate, non-mandatory insurance against losses arising from low probability events, e.g., earthquake, flood and hurricane damage in areas

---

9In the special case of one outcome each in the domain of gains and losses, as in the lottery $\tilde{L}$, cumulative and point transformations of probabilities coincide.

10For the formal definition see Definition 4, below. The axiomatic derivation of the Prelec (1998) function only requires that $\alpha > 0$. However, for mainstream non-EU theories, all of whom postulate the inverse-S shape of $w(p)$, it is required that $0 < \alpha < 1$. We shall refer to the class of Prelec function for which $0 < \alpha < 1$ holds as the standard Prelec functions (Definition 5).
prone to these hazards. The study involved samples of thousands, survey data, econometric analysis and experimental evidence, each gave rise to the same conclusion.11

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. 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’.

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 pinpoints the problem on the demand 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.” 12

This is consistent with the recent evidence from hurricane Katrina. Kunreuther and Pauly (2005, p.65) write: “In the Louisiana parishes affected by Katrina, the percentage of homeowners with flood insurance ranged from 57.7% in St. Bernard’s to 7.3% in Tangipahoa. Only 40% of the residents in Orleans parishes had flood insurance.”

We summarize the main finding in this subsection as stylized fact S2a.

S2a: Many decision makers buy inadequate insurance against low probability events. Whether premiums are actuarially fair, unfair or subsidized, there is a probability below which the take-up of insurance drops dramatically.

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.”

---

11In 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.”

12A 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.
Remark 2: Whilst S2a applies specifically to the insurance behavior for low probability losses, it reflects more general attitudes of decision makers to low probability events. Such events are often simply ignored; see subsection 2.1 below.

1.4. The failure of non-EU models to account for S2a

From Remark 1, ‘standard probability weighting functions’ infinitely overweight infinitesimal probabilities, thus, making such events extremely salient. 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 (see Sections 6, 7). Thus, CP and RDU cannot account for stylized fact S2a. This is a big blow to RDU and CP, because both were proposed precisely to overcome the empirical refutations of EU.

Kahneman and Tversky’s (1979) prospect theory (PT) is capable of incorporating S2a and S2b, but as we note below, it does not give a formal account that lends itself easily to applied work. Furthermore, it allows for the choice of stochastically dominated options (violation of monotonicity) even when such dominance is obvious.

1.5. Insurance for low probability events-II

In contrast to the findings in subsection 1.3, in certain contexts/frames/problems one observes too much insurance for low probability risks, e.g., individuals might oversubscribe to extended warranties or flight insurance rather than life insurance; see Kunreuther and Pauly (2006). Where individuals have strong emotional feelings about an objects, they focus on the salience of the loss irrespective of the low probability; see Rottenstreich and Hsee (2001) and Sunstein (2003). Emotions could be governed by proximity to the source of a loss (e.g., those living within 100 miles of the World Trade Center felt a relatively greater sense of risk), or personal experience or that of friends, or recency of the loss etc.

The evidence shows that individuals who have not claimed on their insurance policies for several years will often cancel their insurance policies; see Kunreuther et al. (1985).

Recent evidence from the choice of deductibles in insurance indicates significant risk aversion for a loss that arises with a probability of about 5% or less; see Sydnor (2010); Cohen and Einav (2007). However, a couple of caveats are in order. First, the deductible choice is typically made by households who are already interested in buying insurance. As far as we can make it, “no-insurance”, was not offered as one of the choices. Second framing issues could have influenced the results.\footnote{Experimental subjects were asked to choose between deductibles of $1000, $500, $250, $100. Most choose the two middle options with almost half the sample choosing the option closes to the average, $500. It is well known in behavioral economics that decision makers can sometimes be unable to choose between \$x, \$y, \$z but when given a choice between \$x, \$y, \$z, they choose \$x or \$z.} Third, a finding of risk aversion for small stakes does not imply that the decision maker will necessarily buy insurance. Nevertheless, in
conjunction with other results in this subsection, these findings are of interest but need further replication in other contexts, problems and frames.

We summarize the findings in this subsection as stylized fact S2b.

\textit{S2b: Many decision makers buy “too much” insurance for low probability events, possibly on account of emotions, claims history, framing/context etc.}

1.6. Combining S2a, S2b: The bimodal perception of risk

Stylized facts S2a, S2b indicate that decision makers either ignore low probability losses or find them to be very salient. There is strong evidence of a \textit{bimodal perception of risks}, see Camerer and Kunreuther (1989) and Schade et al. (2001). A fraction \( \mu \in (0, 1] \) of individuals do not pay attention to losses whose probability falls below a certain threshold (stylized fact S2a), while for the remaining fraction, \( 1 - \mu \), the size of the loss is relatively more salient despite the low probability (stylized fact S2b). McClelland et al. (1993) find strong evidence of a bimodal perception of risk for insurance for low probability losses. Individuals may have a threshold below which they underweight risk and above which they overweight it. Furthermore, individuals may have different thresholds. Across any population of individuals, for any given probability, one would then observe a bimodal perception of risks; see Viscusi (1998). Kunreuther et al. (1988) argue that the bimodal response to low probability events is pervasive in field studies.

Thus, the evidence in S2a and S2b from insurance is consistent with independent evidence from the bimodal perception of risks framework.

1.7. Aims and results of our paper

We argue that a satisfactory theory of insurance should, at the very minimum, be able to address stylized facts S2a, S2b. Furthermore, it should accommodate other findings, as far as possible, on the effect of context and framing on the demand for insurance (see Section 2, below, for more details). As noted above, EU, RDU, CP fail to incorporate S2a. The problems with using PT to explain the facts were noted in subsection 1.4 above. Clearly this is an unsatisfactory state of affairs for the theory of insurance demand.

One must ideally accommodate behavior that is consistent with S2a (and Remark 2) for extreme probabilities, while conforming to S1 for interior probabilities in the interval \([0, 1]\). In order to achieve this, al-Nowaihi and Dhami (2010) in their \textit{composite prospect theory} (CCP) propose a modification to the Prelec function sketched in Figure 1.1. They call their modification as the \textit{composite Prelec weighting function} (CPF) which is sketched in Figure 1.2, below. The CPF is consistent with S1, S2a, is parsimonious yet flexible and has an axiomatic foundation.

In Figure 1.2, for the middle segment, \( p \in [p_1, p_3] \), the CPF is identical to the standard
Figure 1.2: The cumulative probability weighting function, CPF.

Prelec function, and so addresses stylized fact S1. But decision makers heavily underweight very low probabilities in the range $[0, p_1]$ and heavily overweight them in the range $[p_3, 1]$; for evidence on the latter see Kahneman and Tversky’s (1979, p.282-83) and Subsection 8, below. A critical property of the CPF that will prove critical to addressing S2a is the following (contrast this with Remark 1).

Remark 3: For the CPF, $\lim_{p \to 0} w(p)/p = 0$ and $\lim_{p \to 1} \frac{1-w(p)}{1-p} = 0$.

1.7.1. Composite Prospect theory (CCP) and our main results

Under the composite prospect theory (CCP) of al-Nowaihi and Dhami (2010), all decision makers use CP (see Example 2, above) but with one important difference. An exogenous fraction, $\mu$, of individuals replace the standard weighting function (see Figure 1.1, and Remark 1) with the CPF (see Figure 1.2 and Remark 3). Consequently, these individuals respect S1 for interior probabilities but S2a for extreme probabilities. The remaining fraction, $1-\mu$, continue to use the standard probability weighting function (see Remark 1). Consequently, they respect S1 for interior probabilities but can be shown to respect S2b for extreme probabilities. As noted in subsection 1.6, this framework is supported by the bimodal perception of risks framework.

Our first main result is that the fraction $\mu$ of individuals underweight low probabilities sufficiently (in the sense of Remark 3) to ‘not buy’ insurance against such losses (S2a). On the other hand, the fraction $1-\mu$, sufficiently overweights low probabilities (in the sense of Remark 1) that they are keen to buy insurance against such losses (S2b).

Second, we show that CCP can also explain other aspects of the insurance problem that we consider in Section 2. The fraction $\mu$ in CCP is a parameter that depends on the context/frame/problem. Emerging evidence suggests that the take-up of insurance is

---

14 al-Nowaihi and Dhami (2010) also define Composite rank dependent utility (CRDU), which is similar to CCP except that RDU replaces CP. But we argue that CCP is more satisfactory than CRDU.
context dependent and it is influenced by framing effects and emotions. These issues are
discussed below in subsection 2.5. CCP can incorporate these stylized facts through the
variation in $\mu$.

The adverse selection model of insurance makes a prediction for which there is empirical
support. It shows that market failure may arise from the exiting of low-risk types such
that the remaining risk-types are inferior from the viewpoint of insurance firms. In our
third contribution we show that CCP applied to insurance demand also makes a prediction
with similar implications.

We conclude that CCP provides, at the moment, probably the best account for the
demand for insurance for events of all probabilities.

1.8. Structure of the paper

In Section 2 we briefly discuss human behavior from other contexts that supports S2a and
introduce other relevant findings from insurance demand. Section 3 describes the basic
insurance model. Section 4 introduces some essentials of non-linear probability weighting
that are required to apply RDU and CP. Section 5 describes the Prelec probability weight-
ing 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 the composite Prelec probability weighting functions (CPF). Section
9 then combines CPF with, respectively, CP and RDU to form composite prospect theory
(CCP) and composite rank dependent theory (CRDU). Further, it shows that CRDU and
CCP successfully explain the relevant stylized facts. Section 10 concludes. Proofs are
relegated to the Appendix.

2. Other relevant aspects of the insurance problem

In this section we give evidence from non-insurance contexts that supports stylized fact
S2a. We also briefly review how framing and context affects insurance demand and how
this might be accommodated within our framework.

2.1. Other examples of individual response to low probability events (S2a)

In diverse contexts, people ignore very low probability events that could potentially impose
huge losses (as in stylized fact S2a). Many of these losses are self-imposed, on account of
an individual’s own actions, and people choose not to self-insure. It is beyond the scope
of this paper to offer a full discussion but we note some evidence, below. The interested
reader can consult Dhami and al-Nowaihi (2010) for a detailed discussion and the extensive
references to back each of the claims in this subsection.
1. Consider the choice of wearing seat belts in a moving vehicle. There is (at least) a small probability of an accident (self-inflicted) with potentially infinite costs that arise from not wearing seat belts. People were reluctant to wear 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.

2. A user of mobile phones in a moving vehicle faces potentially infinite punishment (e.g., loss of one’s and/or the family’s life) with low probability, in the event of an accident. Evidence indicates that the incidence of driving and talking on mobiles is 40% in the UK, 65% in Finland and 85% in the US.

3. Even as evidence accumulated about the dangers of breast cancer (low probability, high consequence event) free breast cancer examinations were taken up sparingly.

4. Bar-Ilan and Sacerdote (2004) estimate that there are about 260,000 accidents/year in the USA caused by red-light running with implied costs of car repair 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.

5. A fundamental result, the Becker (1968) proposition, states that the most efficient way to deter a crime is to impose the ‘severest possible penalty (to maintain effective deterrence) with the lowest possible probability (to economize on the costs of enforcement)’. However, such punishments are not fully effective; this we call as the Becker paradox. For instance, capital punishment is not effective for all individuals (as in S2a) although it does appear to deter some who would otherwise commit a heinous crime (as in S2b). Historically the severity of punishments has been decreasing. Dhami and al-Nowaihi (2010) examine about 10 potential explanations of the Becker paradox but none of these explains the paradox.

In each of these cases, the evidence is best explained as in the bimodal perception of risks (subsection 1.6) and composite prospect theory, CCP (subsection 1.7.1). In the Nobel prize winning work on prospect theory (PT), Kahneman and Tversky (1979) recognized the contrast between S2a and S2b. Recall our discussion of decision weights $\pi$, in subsection 1.2, Example 1. Kahneman and Tversky (1979) used the particular point transformation

\[15\text{Extensive evidence suggests that the perceived probability of an accident might be even lower than the actual probability because drivers are overconfident of their driving abilities. Some evidence suggest that up to 90 percent of car accidents might be caused by overconfidence.}

\[16\text{In 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).}

9
of probability, $\pi_i = w(p_i)$ (hence, decision weights and probability weights are identical here). They drew $\pi(p)$, as in Figure 2.1, which is undefined at both ends to reflect issues of S2a, S2b.

Kahneman and Tversky’s (1979, p.282-83) summarize the evidence for S2a, S2b as follows. “The sharp drops or apparent discontinuities of $\pi(p)$ at the end-points are consistent with the notion that there is a limit to how small a decision weight can be attached to an event, if it is given any weight at all. A similar quantum of doubt could impose an upper limit on any decision weight that is less than unity...the simplification of prospects can lead the individual to discard events of extremely low probability and to treat events of extremely high probability as if they were certain. Because people are limited in their ability to comprehend and evaluate extreme probabilities, highly unlikely events are either ignored [S2a] or overweighted [S2b], and the difference between high probability and certainty is either neglected or exaggerated. Consequently $\pi(p)$ is not well-behaved near the end-points.”

2.2. The predictions of the insurance model under asymmetric information

In the standard asymmetric information insurance model, the insurance firm does not know the risk type (say, high or low) of the potential buyer of insurance. It is standard in these types of models that there is potential market failure and the following testable hypothesis holds.

$H1$: The market could drive out the low risk types (market failure), leaving only the high risk types active in the market. The high risk types then have higher expected losses and higher claim rates as compared to the population average.


Figure 2.1: Ignorance at the endpoints. Source: Kahneman and Tversky (1979).
2.3. Association of risky activities with demand for insurance

Cutler et al. (2008) find that there is heterogeneity in risk attitudes in various insurance markets. They examine 5 insurance markets: life insurance, acute health insurance, annuities, long-term care insurance, and medicare supplemental insurance (Medigap). They proxy risk tolerance by five measures of attitudes to risk. These are smoking, drinking, job based mortality risk, receipt of preventive health care and use of seat belts.\textsuperscript{17} They then ask if individual insurance purchase is correlated with attitudes to risk. The main finding of interest, from our point of view, is the following.

\textit{F1: Individuals who do not take measures to reduce risk (i.e. buy inadequate supplemental health insurance or do not wear seat belts) are also less likely to buy insurance.}

This evidence supports the bimodal perception of risks, prospect theory (PT) and composite prospect theory (CCP). Individuals who do not wear seat belts seem to be ignoring the very low probability of a potentially fatal accident. As in S2a, we might also expect them to ‘not insure’ for very low probability losses, which is indeed the case that is consistent with the evidence.

2.4. Context dependence and framing effects on insurance demand

Standard economic theory assumes context invariance of risk preference, and, hence, context invariance of insurance demand. Recent research suggests evidence to the contrary. For instance, Barseghyan et al. (2010) find that risk preferences depend on the kind of insurance, e.g., auto-collision, auto comprehensive, and home all perils.

In Hershey et al. (1982) the framing of lotteries as ‘gamble’ or ‘insurance’ affects individual choice. Johnson et al. (1993) find that the framing of coinsurance as ‘deductible’ or ‘rebate’ affects choices. Context dependent risk preferences are also found by Wolf and Pohlman (1983) for dealers in US government securities.\textsuperscript{18} The main finding of this subsection is summarized below.

\textit{F2: The amount of insurance purchased is dependent on the context/frame/problem.}

A desirable model of insurance should also be able to take account of F1, F2.

\textsuperscript{17} Cutler et al (1998) argue that these five variables capture individual risk aversion. Our view is that this interpretation is too strong. We simply refer to these as attitudes to risk. For instance, smoking and drinking are addictions and a large branch of behavioral economics addresses these as self control problems. We are also skeptical about job based mortality risk as a measure of individual risk aversion. However, the remaining two measures are quite suitable.

\textsuperscript{18} Consider another example of framing from Kunreuther and Pauly (2006, p.96). “If one asks an individual whether she would pay $140 to protect herself against an event with probability $p = 1/100$ and [loss] $L = 10,000$, many people will say “no”. On the other hand, if one frames the same problem as purchasing an insurance policy which costs $140, a much higher percentage of people will say “yes”.
2.5. Determinants of $\mu$

The bimodal perception of risks framework and PT do not specify the size of the fraction $\mu$ that ignores very low probability events. In the insurance context, $F_2$ in subsection 2.4 and $S_{2b}$ in subsection 1.5, suggest that $\mu$ is likely to depend on the context and the framing of the problem. We believe that the salience of low-probability large losses (and so the size of $\mu$) can be influenced by the media, family, friends, and public policy. Consider two examples from subsection 2.1. The take-up of free breast cancer examination rose greatly after the greatly publicized mastectomies of Betty Ford and Happy Rockefeller. Vivid public warnings of the fatal consequences of speeding (viewed by some as low probability events) can have a similar effect. The discussion in subsections 1.5, 2.3, 2.4 indicates that the size of $\mu$ can also be influenced by emotions, experience, time available to make a decision, bounded rationality, framing, incentive effects and so on.\(^{19}\)

We do not specify the precise source of $\mu$. Rather, we shall simply take $\mu$ as a given parameter, hence, our framework is compatible with all the potential explanations, above.

2.6. Are these low probabilities economically relevant?

In subsections 8.1, 8.2 we fit two composite Prelec probability weighting functions, CPF (see Figure 1.2), to two separate data sets from Kunreuther et al. (1978). The intervals of low probabilities that are underweighted by the CPF (the range $(0, p_1]$ in Figure 1.2) are, respectively, $(0, 0.195]$ and $(0, 0.006]$. The upper end of the first interval might seem too high relative to one’s gut feeling but this is what the evidence shows up. Furthermore, the decision maker might still wish to buy insurance even if he/she underweights these probabilities.

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.\(^{20}\) We allow departures from the actuarially fair condition. We do so in a simple way by setting the insurance premium rate, $r$, to be

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

\(^{19}\)See also chapter 4 titled ‘Anomalies on the demand side’ in Kunreuther and Pauly (2006).

\(^{20}\)The 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.
Thus, \( \theta = 0 \) corresponds to the actuarially fair condition, and \( \theta > 0, \theta < 0 \) to respectively the actuarially unfair and ‘over-fair’ conditions. Hence, the decision maker’s wealth is:

\[
\begin{cases}
W - rC - f & \text{with probability } p \\
W - rC - f - (L - C) & \text{with probability } 1 - p
\end{cases}
\]

Let \( U_I(C) \) be the utility of the decision maker from an amount of coverage, \( C > 0 \) and \( U_{NI} \) the utility from ‘no insurance’. Let \( C^* \) be the optimal level of coverage, conditional on buying insurance. The decision maker must, in addition, satisfy the participation constraint,

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

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 transformations of the cumulative probability distribution. In this section, we introduce the 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 \( w(p) : [0, 1] \rightarrow [0, 1] \).

**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 \( \lim_{p \to 0} \frac{w(p)}{p} = \infty \), and (b) infinitely-underweights near-one probabilities if \( \lim_{p \to 1} \frac{1-w(p)}{1-p} = \infty \).

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

Evidence strongly suggest that decision makers exhibit an inverse S-shaped probability weighting function over outcomes (recall stylized fact S1, Figure 1.1 and Remark 1, above).

**Remark 4** (Standard probability weighting functions): A large number of probability weighting functions have been proposed, e.g., those by Gonzalez and Wu (1999) and Lattimore, Baker and Witte (1992), Tversky and Kahneman (1992). They all infinitely overweight infinitesimal probabilities and infinitely underweight near-one probabilities (Definition 2). We shall call these as the standard probability weighting functions. and these are the ones used in RDU and CP.
5. Prelec’s probability weighting function

The Prelec (1998) probability weighting function is parsimonious and axiomatically founded. However, it can be made consistent with either S1 or S2, but not both.\footnote{While we have chosen the Prelec function for illustration, these comments apply to almost all weighting functions used in RDU and CP.}

**Definition 4** (Prelec, 1998): By the Prelec function we mean the probability weighting function \(w(p) : [0,1] \rightarrow [0,1]\) given by \(w(0) = 0, w(1) = 1\) and

\[
\begin{align*}
w(p) &= e^{-\alpha(-\ln p)\beta}, & 0 < p \leq 1, & \alpha > 0, \beta > 0. 
\end{align*}
\]

(5.1)

We make a distinction between the Prelec function and the standard Prelec function; Prelec (1998) prefers the latter for reasons we give below.

**Definition 5** (Standard Prelec function): In the standard Prelec function, used in RDU and CP, the value of \(\alpha\) in (5.1) is restricted to \(0 < \alpha < 1\).

**Proposition 2**: The Prelec function (Definition 4) is a probability weighting function in the sense of Definition 1.

We now look at the roles of \(\alpha, \beta\) in the Prelec function.\footnote{For the full set of possibilities for the Prelec function see al-Nowaihi and Dhami (2010).}

1. (Role of \(\alpha\)) The parameter \(\alpha\) 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. In this case, it is inverse-S shaped, as in \(w(p) = e^{-(-\ln p)^{\frac{1}{\alpha}}} (\alpha = 0.5, \beta = 1)\), which is sketched as the thick curve in Figure 5.1. The converse holds if \(\alpha > 1\), in which case the Prelec function is S shaped. An example is the curve \(w(p) = e^{(-(-\ln p)^{2}} (\alpha = 2, \beta = 1)\), sketched in Figure 5.1 as the light curve (the straight line in Figure 5.1 is the \(45^\circ\) line).

2. (Role of \(\beta\)) Between the region of strict convexity \((w'' > 0)\) and the region of strict concavity \((w'' < 0)\), for each of the cases in 5.1, there is a point of inflexion \((w'' = 0)\). The parameter \(\beta\) controls the location of the inflexion point relative to the \(45^\circ\) line. Thus, for \(\beta = 1\), the point of inflexion is at \(p = e^{-1}\) and lies on the \(45^\circ\) line, as in Figure 5.1. However, if \(\beta < 1\), then the point of inflexion lies above the \(45^\circ\) line, as in \(w(p) = e^{-\frac{1}{2}(-\ln p)^{2}} (\alpha = 2, \beta = \frac{1}{2})\). For this example, the fixed point, \(w(p^*) = p^*\), is at \(p^* \simeq 0.14\) but the point of inflexion, \(w''(\tilde{p}) = 0\), is at \(\tilde{p} \simeq 0.20\).

In Figure 5.1, when \(\alpha < 1\) (inverse-S shape), \(w(p)\) becomes very steeply sloped near \(p = 0\) and \(p = 1\). By contrast, these slopes are very gentle for \(\alpha > 1\).
Figure 5.1: Plots of $w(p) = e^{-\frac{1}{2}(-\ln p)^2}$ and $w(p) = e^{-(\ln p)^2}$.

**Proposition 3**: (a) For $\alpha < 1$ the Prelec (1998) function: (i) infinitely-overweights infinitesimal probabilities, i.e., $\lim_{p \to 0} \frac{w(p)}{p} = \infty$, and (ii) infinitely underweights near-one probabilities, i.e., $\lim_{p \to 1} \frac{1-w(p)}{1-p} = \infty$.

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

According to Prelec (1998, p.505), the infinite limits in Proposition 3a capture the qualitative change as we move from certainty to probability and from impossibility to improbability. On the other hand, they contradict stylized fact that many people ignore events of very low probability and treat very high probability events as certain (see S2a and subsection 2.1). 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 in S2a.

These specific problems are avoided for $\alpha > 1$. However, for $\alpha > 1$, the Prelec function is *S-shaped*, see Figure 5.1. This is in conflict with the empirical evidence on stylized fact S1, which indicates an *inverse-S shape* for probabilities bounded away from the end points of the interval $[0,1]$. Hence, the two cases, $\alpha < 1$ or $\alpha > 1$, by themselves, are unable to simultaneously address stylized facts S1 and S2a.

**6. Rank dependent utility (RDU) and insurance**

We now model the behavior of an individual using rank dependent utility theory (RDU); see Example 1. Consider a decision maker with utility of wealth, $u$, 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 above\(^{23}\) by (say) $u'_{\text{max}}$.

\[^{23}\]The boundedness of $u'$ is needed for Proposition 4b. 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 $x^\gamma$, $0 < \gamma < 1$. By contrast, the boundedness of $u'$ is not a requirement in CP or CCP.
If the decision maker buys insurance, then using Example 1 in the introduction, the utility under RDU is:

\[
U_I (C) = [1 - w (1 - p)] u (W - rC - f - L + C) + w (1 - p) u (W - rC - f) .
\] (6.1)

**Definition 6** (Major and minor loss): Let \( L_M = rC + f + L - C \) be the major loss and \( L_m = rC + f < L_H \) be the minor losses to the decision maker. We then rewrite (6.1) as:

\[
U_I (C) = [1 - w (1 - p)] u (W - L_M) + w (1 - p) u (W - L_m) .
\] (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, \( C = L \), (6.1) gives:

\[
U_I (L) = u (W - rL - f) .
\] (6.3)

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

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

The following **participation constraint** must be satisfied to buy insurance coverage \( C^* \):

\[
U_{NI} \leq U_I (C^*)
\] (6.5)

**Proposition 4**: 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 \( \overline{L} F (p) \), where \( \overline{L} = pL \) and

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

(b) If the probability weighting function infinitely-underweights near-one probabilities (Definition 2b) then, for a given expected loss, \( \overline{L} \), 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, \( \overline{L} \), a decision maker will not insure, for all sufficiently small probabilities.

To see the intuition, first consider Proposition 4(b). 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 particular, for the Prelec function with \( \alpha < 1 \) (see Definition 4, Proposition 3(aii), Remark 4). In this case (see Figures 1.1, 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 (6.2). As $p \to 0$, $1 - p \to 1$ and, hence, as required in Proposition 4(b), $w(1 - p)$ underweights $1 - p$. This reduces the relative salience of the second term in (6.2), relative to the first term, hence, highlighting the salience of the major loss, $L_M$. This makes insurance even more attractive under RDU than under EU (which was already too high, given the evidence).

The reverse occurs in Proposition 4(c), as the reader can readily check. Here the probability weighting function zero-overweights near-one probabilities (Definition 3b), which is true for the composite Prelec function, CPF, that is used in composite prospect theory, CCP (see Remark 3).

In this case, as $p \to 0$, $1 - p \to 1$ and, hence, as required in Proposition 4(c), $w(1 - p)$ overweights $1 - p$. By reducing the salience of the first term relative to the second in (6.2), this makes insurance against very low probability events unattractive, in conformity with the evidence. Proposition 4(c) is further discussed in section 8, below.

It is of interest to get a feel for how restrictive this participation constraint in Proposition 4(a) is. Example (3), below, suggests that it is a weak restriction.

Example 3: To check the restrictiveness of the participation constraint we use the result in Proposition 4(a), i.e., the sufficient condition $f \leq LF(p)$ for the participation constraint to hold. The first row of the Table, below, gives losses, $L$ (in dollars, say), from 10 to 10^7, with corresponding probabilities, $p$, in row 2, ranging from 10^{-1} to 10^{-7}. Hence, the expected loss in each case is $L = 1$, 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. Rows 4, 5 can be used to check the restrictiveness of the condition in Proposition 4(a).

---

24 This is also the case for the Prelec function with $\alpha > 1$ (Definition 4 and Proposition 3(bii)). However, the first segment of the CPF is the same as the Prelec function for the case $\alpha > 1$ (see section 8, below).

25 This methodology is motivated by the empirical evidence in Kunreuther et al. (1978) who vary $p, L$ while keeping $L$ constant.

26 $\theta = 1$ offers an even more stringent test of the restrictiveness of the participation constraint, as compared to lower values of $\theta$. 

17
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 4(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.

7. Cumulative prospect theory (CP) and insurance

The essential features of CP are outlined in Example 2. Two utility functions, \( v(y) \), that satisfy the properties listed in Example 2 are \( v(x) = 1 - e^{-x} \) 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:\(^{27}\)

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

Based on experimental evidence, Kahneman and Tversky (1979) suggest \( \gamma = 0.88 \), \( \lambda = 2.25 \).

**Example 4** (Generalizing Example 2) Consider the general prospect

\[
\tilde{L} = (y_{-m}; p_{-m}; \ldots; y_{-1}; p_{-1}; y_0; p_0; y_1; p_1; \ldots; y_n; p_n),
\]

that has \( m \) outcomes in the domain of loss and \( n \) outcomes in the domain of gains. In this case, the decision weights for gains are

\[
\pi_i = w\left(\sum_{j=i}^n p_j\right) - w\left(\sum_{j=i+1}^n p_j\right), \quad i = -1, \ldots, -m
\]

and for losses they are

\[
\pi_j = w\left(\sum_{i=-m}^j p_i\right) - w\left(\sum_{i=-m-1}^j p_i\right), \quad j = 1, \ldots, n. \quad ^{28}\]

With the decision weights constructed in this manner, a decision maker evaluates the lottery \( \tilde{L} \) as

\[
V(\tilde{L}) = \sum_{i=-m}^n \pi_i v(y_i), \quad \text{where } v(y) \text{ is defined in (7.1)}.
\]

**Remark 5** : The reader should note well that in CP and in RDU, \( w \) is a standard probability weighting function (see Remark 4). An example is the inverse S-shaped Prelec function for the case \( \alpha < 1 \) (see Definition 4) that satisfies stylized fact S1. This has the important implication, from Proposition 3(a), that small probabilities are infinitely-overweighted,

\(^{27}\)For an axiomatic derivation of this value function, see al-Nowaihi et al. (2008).

\(^{28}\)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), an assumption we too make.
\[ \lim_{p \to 0} \frac{w(p)}{p} = \infty. \quad \text{In particular, this rules out the fact that small probabilities could be infinitely-underweighted, i.e., } \lim_{p \to 0} \frac{w(p)}{p} = 0, \text{ as in (i) the S-shaped Prelec function for } \alpha > 1, \text{ which violates S1 (see Proposition 3(b)), and (ii) the composite Prelec function (CPF) (see Remark 3 and Proposition 8, below).} \]

### 7.1. The insurance decision of an individual who uses CP

Consider a decision maker who follows CP and faces the insurance problem of Section 3. Recall Example 2 in the introduction. The reference point is taken to be the status-quo wealth, \( W \). With probability \( 1 - p \), wealth relative to the reference wealth is

\[ (W - rC - f) - W = -rC - f < 0. \tag{7.2} \]

With the complementary probability \( p \), wealth relative to the reference point is

\[ (W - rC - f - L + C) - W = -rC - f - (L - C) \leq -rC - f < 0. \tag{7.3} \]

Notice from (7.2), (7.3) that the decision maker is in the ‘domain of loss’ in both states. Using Example 4, the decision maker’s value function under CP when some level of insurance coverage \( C \in [0, L] \) is purchased is given by

\[ V_I(C) = w(p) v(-rC - f - (L - C)) + [1 - w(p)] v(-rC - f). \tag{7.4} \]

Using the definition of a major and a minor loss (Definition 6) rewrite (7.4) as:

\[ V_I(C) = w(p) v(-L_M) + [1 - w(p)] v(-L_m) \tag{7.5} \]

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, \( C = L \), (7.4) gives

\[ V_I(L) = v(-rL - f). \tag{7.6} \]

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 \( v(0) = 0 \)):

\[ V_{NI} = w(p) v(-L). \tag{7.7} \]

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:

\[ V_{NI} \leq V_I(C^*). \tag{7.8} \]
Proposition 5: 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.\(^{29}\) This holds even in the presence of actuarially unfair premiums and a fixed cost of insurance.

(b) For Prelec’s probability weighting function (Definition 4), with $\alpha < 1$, for the value function (7.1) and for a given expected loss, $\overline{L}$, the participation constraint (7.8) is satisfied for all sufficiently small probabilities. A sufficient condition for the satisfaction of the participation constraint is that $\beta < \overline{L} F(p)$ where

$$F(p) = \frac{e^{-\frac{\beta}{\gamma}(-\ln p)^\alpha}}{p} - (1 + \theta), \ 0 < \alpha < 1, \ \beta > 0, \ \gamma > 0.$$ \hspace{1cm} (7.9)

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

The reason that part (a) holds is very simple. Under CP, the value function is strictly convex for losses, hence, a decision maker will always insure fully, if he insures at all. This is, of course, at variance with the evidence on partial coverage. For the intuition for parts (b), (c), the reader should bear in mind Remark 5.

The intuition behind part (b) is as follows. For $\alpha < 1$ (which is the case under RDU and CP) the Prelec function infinitely-overweights infinitesimal probabilities (Definitions 2a, 4 and Proposition 3(ai)). Now return to (7.5). As $p \to 0$, and as required by Proposition 5(b), $w(p)$ increasingly overweights $p$.\(^{30}\) This increases the relative salience of the first term (i.e. major losses) in (7.5) relative to the second term. 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). In this case, as $p \to 0$ and under the conditions of Proposition 5(c)\(^{31}\) the relative salience of the first term (major loss) declines at the expense of the second term in (7.5). This makes insurance against very low probability events unattractive, in conformity with the evidence.\(^{32}\)

By Proposition 5(a), a decision maker will insure fully against any loss, provided the participation constraint (7.8) is satisfied, even with fixed costs of insurance and an actuarially unfair premium. By Proposition 5(b), for Prelec’s probability weighting function

\(^{29}\)Which of the two cases holds, therefore, depends on the parameters of the model and so, we require simulations. See Example 5, below.

\(^{30}\)In the sense that $w(p)/p$ increases as $p \to 0$.

\(^{31}\)These conditions hold true of a Prelec function with $\alpha > 1$ and the composite Prelec function, CPF (see Remark 3).

\(^{32}\)Proposition 5(c) will help us show that composite prospect theory (CCP) explains the evidence for the take-up of insurance for low probability events; Section 8, below.
(Definition 4), for the value function (7.1) and for a given expected loss, the participation constraint (7.8) is satisfied for all sufficiently small probabilities. Example 5, below, shows that this participation constraint imposes only a very weak restriction.

**Example 5**: The first row of the following Table gives losses (in dollars, say) from 10 to $10^7$, with corresponding probabilities (row 2) ranging from $10^{-1}$ to $10^{-7}$; so that the expected loss in each case is $\bar{L} = 1.33$ In row 3 are the corresponding values of $e^{\frac{2}{\alpha}(\ln p)^{\beta}}$ (see (7.9)), 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 (7.9)) for the high profit rate of 100% ($\theta = 1$) for the insurance firm, so that $F(p) = e^{\frac{2}{\alpha}(\ln p)^{\beta}} - 2.34$ Row 5 gives the (large) upper bound on fixed costs as a percentage of expected losses, which, if it holds, ensures that the results in Proposition 5(a),(b) go through.

<table>
<thead>
<tr>
<th>$p$</th>
<th>$10^{-1}$</th>
<th>$10^{-2}$</th>
<th>$10^{-3}$</th>
<th>$10^{-4}$</th>
<th>$10^{-5}$</th>
<th>$10^{-6}$</th>
<th>$10^{-7}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$e^{\frac{2}{\alpha}(\ln p)^{\beta}}$</td>
<td>1.4169</td>
<td>4.6589</td>
<td>18.48</td>
<td>81.342</td>
<td>383.83</td>
<td>1906.3</td>
<td>9852.3</td>
</tr>
<tr>
<td>$F(p)$</td>
<td>-0.5831</td>
<td>2.6589</td>
<td>16.48</td>
<td>79.342</td>
<td>381.83</td>
<td>1904.3</td>
<td>9848.3</td>
</tr>
<tr>
<td>$\frac{F(p)}{L} \times 100$</td>
<td>-58.31</td>
<td>265.89</td>
<td>1648</td>
<td>7934.2</td>
<td>38183</td>
<td>190430</td>
<td>984830</td>
</tr>
</tbody>
</table>

It is obvious that the participation constraint is hardly restrictive for low probabilities. Thus, from Proposition 5(a),(b), we see that CP is likely to lead to misleading results, because it predicts too much insurance.

8. The composite Prelec function, CPF

Our aim is to introduce *composite Prelec functions* (CPF), which when combined with either RDU or CP, can explain S2a, S2b (see introduction), F1, and F2 (see Section 2). We begin, in sections 8.1, 8.2, by providing two numerical examples of CPF that are motivated by the empirical evidence from Kunreuther (1978). This is followed, in subsection 8.3, by a more formal treatment of CPF.

8.1. First example: The farm experiments in Kunreuther (1978)

In their “farm” experiments Kunreuther et al. (1978, ch.7) report that the take-up of actuarially fair insurance declines if the probability of the loss (keeping the expected loss

---

33 As noted earlier, this conforms to the practice in Kunreuther et al (1978).

34 As noted in the earlier example, a higher value of $\theta$ provides a more stringent test of the restrictiveness of the participation constraint.
constant) goes below 0.05. To capture this finding, we postulate a CPF composed of segments from three Prelec functions, and given by

\[
w(p) = \begin{cases} 
  e^{-0.19286(-\ln p)^2}, & \text{i.e., } \alpha = 2, \beta = 0.19286, \text{ if } 0 < p < 0.05 \\
  e^{-(\ln p)^{1/2}}, & \text{i.e., } \alpha = 0.5, \beta = 1, \text{ if } 0.05 \leq p \leq 0.95 \\
  e^{-86.081(-\ln p)^2}, & \text{i.e., } \alpha = 2, \beta = 86.081, \text{ if } 0.95 < p \leq 1
\end{cases} \quad (8.1)
\]

The CPF in (8.1) is plotted in Figure 8.1. For \(0 \leq p < 0.05\), the CPF is identical to the S-shaped Prelec function, \(e^{-\beta_0(-\ln p)\alpha_0}\), with \(\alpha_0 = 2, \beta_0 = 0.19286\). For \(0.05 \leq p \leq 0.95\), the CPF is identical to the inverse-S shaped Prelec function with \(\alpha = 0.5, \beta = 1\). For \(0.95 < p \leq 1\), the CPF is identical to the S-shaped Prelec function, \(e^{-\beta_1(-\ln p)\alpha_1}\), with \(\alpha_1 = 2, \beta_1 = 86.081\).

The first and the third segments of the CPF correspond to the light Prelec curve in Figure 5.1 (with \(\alpha > 1\)), while the second segment corresponds to the thick Prelec curve in Figure 5.1 (with \(\alpha < 1\)). \(\beta_0\) is chosen to make \(w(p)\) continuous at \(p = 0.05\), while \(\beta_1\) is chosen to make \(w(p)\) continuous at \(p = 0.95\).

**Remark 6** (Fixed points): This CPF (8.1) 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 7**: (Underweighting and overweighting of probabilities): The CPF in Figure 8.1 overweights ‘low’ probabilities, in the range \(0.0055993 < p < e^{-1}\) and underweights ‘high’ probabilities, in the range \(e^{-1} < p < 0.98845\). Figures 8.2 and 8.3, below, respectively, magnify the regions near 0 and near 1. The CPF underweights ‘very low’ probabilities, in the range \(0 < p < 0.0055993\) and overweightes ‘very high’ probabilities, in the range \(0.98845 < p < 1\). 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). The CPF is also nearly flat near 1. We will show that these two segments, i.e., \( p \in (0, 0.0055993) \cup (0.98845, 1) \) are able to address S2a.

![Figure 8.2: Behaviour of Figure 8.1 near 0.](image1.png)

![Figure 8.3: Behaviour of Figure 8.1 near 1.](image2.png)

**Remark 8**: For other data sets (see e.g., the urn experiments, below), one need not obtain the same, or similar, cut-off points in (8.1). The CPF allows for such flexibility.

**Remark 9**: In representing the data from Kunreuther et al. (1978) we have chosen the values of \( \alpha \) to give the correct shape of the CPF. We then chose the values of \( \beta \) to ensure continuity across the three segments of the Prelec function. The converse, i.e., choosing \( \beta \) values first and then fixing the \( \alpha \) values to ensure continuity is less attractive because the \( \alpha \) values control the degree of concavity/convexity of the Prelec function and, so, must ensure that the CPF is of the “correct shape” as in Figure 1.2.

### 8.2. Second example: The urn experiment in Kunreuther (1978)

In their “urn experiments”, Kunreuther et al. (1978, chapter 7) report that 80% of subjects facing actuarially fair premiums took up insurance against a loss with probability 0.25.
But the take-up of insurance declined when the probability of the loss declined, keeping the expected loss constant. At a probability of 0.001, only 20% took up insurance. These results motivate the following CPF,

\[
w(p) = \begin{cases} 
  e^{-0.61266(-\ln p)^2}, & \text{i.e., } \alpha = 2, \beta = 0.61266, \text{ if } 0 \leq p < 0.25, \\
  e^{-(-\ln p)^{1/2}}, & \text{i.e., } \alpha = 0.5, \beta = 1, \text{ if } 0.25 \leq p \leq 0.75, \\
  e^{-6.4808(-\ln p)^2}, & \text{i.e., } \alpha = 2, \beta = 6.4808, \text{ if } 0.75 < p \leq 1.
\end{cases} \tag{8.2}
\]

The plot of the CPF in (8.2) and it’s properties are similar to the case in subsection 8.1. The interested reader can find the details and discussion in al-Nowaihi and Dhami (2010).

### 8.3. A more formal treatment of the CPF

We give a minimal formal treatment of the CPF here. The full axiomatic derivation and details can be found in al-Nowaihi and Dhami (2010). The upper cutoff points for the first segment of the CPF in (8.1), (8.2) are respectively, at probabilities 0.25 and 0.05. Denote this cutoff point as \( p \). Similarly, the upper cutoff point for the second segment of the CPF in Figures 1.2 and 8.1 are respectively, at probabilities 0.75 and 0.95. Denote this cutoff point as \( \overline{p} \) (see Figure 1.2 for the points, \( p, \overline{p} \)). Now define,

\[
p = e^{-\left(\frac{\beta}{\alpha_0}\right)^{\frac{1}{n_0-\alpha}}}, \quad \overline{p} = e^{-\left(\frac{\beta}{\alpha_1}\right)^{\frac{1}{n_1-\alpha}}}.
\]

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

**Definition 7** (Composite Prelec weighting function, CPF; al-Nowaihi and Dhami, 2010a): By the composite Prelec weighting function we mean the probability weighting function \( w: [0, 1] \to [0, 1] \) given by

\[
w(p) = \begin{cases} 
  0 & \text{if } p = 0 \\
  e^{-\beta_0(-\ln p)^{\alpha_0}} & \text{if } 0 < p \leq \overline{p} \\
  e^{-\beta(-\ln p)^{\alpha}} & \text{if } \overline{p} < p \leq \overline{p} \\
  e^{-\beta_1(-\ln p)^{\alpha_1}} & \text{if } \overline{p} < p \leq 1
\end{cases} \tag{8.4}
\]

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

\[
0 < \alpha < 1, \ \beta > 0; \ \alpha_0 > 1, \ \beta_0 > 0; \ \alpha_1 > 1, \ \beta_1 > 0, \ \beta_0 < 1/\beta^{\alpha_0-1}, \ \beta_1 > 1/\beta^{\alpha_1-1}. \tag{8.5}
\]

More generally, al-Nowaihi and Dhami (2010) propose a definition of a CPF that allows for \( n \geq 1 \) segments. For \( n = 1 \), the CPF reduces to the Prelec function. For \( n = 3 \) we get the three piece CPF of Definition 7, illustrated in Figure 1.2 of the introduction.

**Proposition 6**: The composite Prelec function (Definition 7) 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 Dhami (2006)). The restrictions $\beta_0 < 1/\beta^{\alpha_0 - 1}$ and $\beta_1 > 1/\beta^{\alpha_1 - 1}$ ensure that the interval $(\underline{p}, \overline{p})$ is not empty. The interval limits are chosen so that the CPF in (8.4) is continuous across them. Define $p_1, p_2, p_3$ (see Figure 1.2) as.

\[ p_1 = e^{-(\frac{1}{p_0})^{\alpha_0 - 1}}, \quad p_2 = e^{-(\frac{1}{p_1})^{\alpha_1 - 1}}, \quad p_3 = e^{-(\frac{1}{p_1})^{\alpha_1 - 1}}. \]  

(8.6)

**Proposition 7**: (a) $p_1 < p < p_2 < p_3$. (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 6, the CPF in (8.4), (8.5) is a probability weighting function in the sense of Definition 1. By Proposition 7, 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); see subsection 2.1, above.

The restrictions $\alpha_0 > 1$ and $\alpha_1 > 1$ in (8.5) ensure that a CPF has the following properties, that will help explain human behavior for extremely low probability events (see Remark 5).

**Proposition 8**: The CPF (8.4):

(a) zero-underweights infinitesimal probabilities, i.e., \( \lim_{p \to 0} \frac{w(p)}{p} = 0 \) (Definition 3a),

(b) zero-overweights near-one probabilities, i.e., \( \lim_{p \to 1} \frac{1-w(p)}{1-p} = 0 \) (Definition 3b).

9. Explaining the stylized facts of insurance

We now outline al-Nowaihi and Dhami’s (2010) composite cumulative prospect theory (CCP) and composite rank dependent theory (CRDU) and show how these theories can explain the stylized facts of insurance S2a, S2b and are also capable of incorporating the findings F1, F2 in Section 2.


Recall from Remark 5 that CP uses standard probability weighting functions. We offer the following definition of composite cumulative prospect theory, CCP (see also subsection 1.7.1 in the introduction).
Definition 8 (CCP; al-Nowaihi and Dhami, 2010): In composite prospect theory, CCP, there are two types of individuals. An exogenous fraction $1 - \mu$ of the individuals use cumulative prospect theory, CP (Examples 2, 4 and Remark 5). Consequently, these individuals respect stylized fact S1. The remaining fraction, $\mu$, of individuals replace the standard weighting function (see Remark 4) for which $\lim_{p \to 0} w(p) = \infty$ (Proposition 3(a)) with the composite Prelec function, CPF (see Section 8) for which $\lim_{p \to 0} w(p) = 0$ (Proposition 8(a)). Consequently, these individuals respect S1 for interior probabilities but (as required for S2a) ignore near zero probabilities and treat near one probabilities as certainty. Section 2.5 shows how the finding F2 has a bearing on the size of $\mu$.

Proposition 9: Under composite prospect theory (CCP), for a given expected loss, the fraction $\mu$ of decision makers will not insure, for all sufficiently small probabilities. The remaining fraction $1 - \mu$ of decision makers use CP, hence, their behavior is given by Proposition 5(a),(b). In conjunction with Example 5, which shows that the participation constraint is very mild, the fraction $1 - \mu$ of decision makers will choose to insure fully for a given expected loss, for all sufficiently small probabilities.

From Proposition 9, CCP predicts two kinds of insurance outcomes. Some individuals (fraction $\mu$) will not buy any insurance for losses of extremely small probability (stylized fact S2a). Other individuals (fraction $1 - \mu$) will buy insurance for such events (stylized fact S2b). In different contexts and frames (which will affect $\mu$; see Section 2.5) one expects the extent of the take-up of insurance to be varied.

The predictions of the insurance model under CCP are similar in one respect with those of the insurance model under asymmetric information (see subsection 2.2 and Hypothesis H1). Suppose that a heterogenous population of individuals differ with respect to some potential loss that occurs with a probability that lies over some interval $[0, \hat{p}]$ where $0 < \hat{p} \leq 1$. Suppose also that these individuals can buy insurance if they wish to. The fraction $\mu$ of individuals will ignore low probability risks and not seek insurance. The remaining individuals in the market (fraction $1 - \mu$) are, from the insurance firm’s point of view, higher risk individuals. Thus in equilibrium, we would observe the low risks dropping out and the market being dominated by higher risks as is the case under adverse selection. Furthermore, those who insure will have a higher claim rate than the overall population average of risks. Thus, H1 need not hold under an asymmetric information model alone and confirmation of H1 is not unique confirmation of such a model.

We now provide some illustrative examples motivated by Kunreuther et al’s (1978) urn experiments; see subsection 8.2 above.

Example 6 (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., $r = p$. We
assume that the utility function is given in (7.1) with experimental values suggested by Kahneman and Tversky (1979) to be \( \gamma = 0.88 \), and \( \lambda = 2.25 \). 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 (7.8), we need to check, in each case, the following condition that ensures full insurance (we have used the actuarially fair condition \( r = p \)):

\[
 w(p) v(-L) \leq v(-pL)
\]

Substituting (7.1) this becomes, \(-\lambda w(p)L^\gamma \leq -\lambda p^\gamma L^\gamma \). Using \( \gamma = 0.88 \) we get

\[
\frac{w(p)}{p^{0.88}} \geq 1.
\]

(a) Decision maker uses CP: Under CP, suppose that the decision maker uses the Prelec probability weighting function, \( w(p) = e^{-\beta(-\ln p)^\alpha} \) with \( \beta = 1 \) and \( \alpha = 0.50 \). In this case (9.2) requires that

\[
\frac{e^{-(-\ln 0.001)^{0.50}}}{(0.001)^{0.88}} \geq 1 \iff 31.518 \geq 1, \text{ which is true.}
\]

So, a decision maker using CP will fully insure. But, as noted in section 8.2, Kunreuther et al’s (1978) data shows that only 20% of the decision makers insure in this case.

(b) Decision maker uses CCP: Decision makers who belong to the fraction \( 1-\mu \) use CP and so, as in (a), will fully insure. Now consider those who belong to the fraction \( \mu \). Assume that they use the composite Prelec function given in (8.2). Since \( p = 0.001 \in (0, 0.25) \), (8.2) gives \( w(0.001) = e^{-0.61266(-\ln 0.001)^2} \). In this case (9.2) requires that

\[
\frac{e^{-0.61266(-\ln 0.001)^2}}{(0.001)^{0.88}} \geq 1 \iff 8.7838 \times 10^{-11} \geq 1, \text{ which is not true.}
\]

Hence, a decision maker under CCP who belongs to the fraction \( \mu \) of individuals who respect S2a will not insure. Thus, Kunreuther et al’s (1978) data can be explained by CCP with \( \mu = 0.8 \) and \( 1-\mu = 0.2 \).

Example 7 (Urn experiment): We continue to use the set-up of Example 6. However, let the probability of the loss be \( p = 0.25 \) (instead of \( p = 0.001 \)). Since \( p = 0.25 \in [0.25, 0.75] \), (8.2) gives \( w(0.25) = e^{-(\ln 0.25)^{0.50}} \). For this probability, CP and CCP make identical predictions because over the relevant range they share the same probability weighting function. In this case (9.2) requires that

\[
\frac{e^{-(\ln 0.25)^{0.50}}}{(0.25)^{0.88}} \geq 1 \iff 1.0434 \geq 1, \text{ which is true.}
\]
Hence, such a decision maker will insure fully against any loss whose probability of occurrence \( p \geq 0.25 \). Kunreuther et al’s (1978) data shows that 80% of the experimental subjects took up insurance in this case. Hence, for losses whose probability is bounded well away from the end-points, the predictions of, both, CP and CCP are in close (but not perfect) conformity with the evidence.

In conjunction, Examples 6 and 7 illustrate how CCP can account well for the evidence for events of all probabilities while CP’s predictions for low probability events are incorrect. Furthermore, CCP explains S2a, S2b, while CP (like RDU) only explains S2b. Thus, CCP explains everything that CP or RDU can, but the converse is false.

We would like to preempt a potential criticism of CCP. CCP has more parameters than CP, so one might argue that it is not a surprise that it performs better. It is often the case in science that if a theory B explains everything that theory A does and more, then theory B has more parameters. This potential criticism only makes sense if one could come up with a decision theory that explains everything that CCP does and more, yet has a smaller number of parameters. Clearly EU, RDU, and CP are not potential candidates for this exercise. Indeed for situations where S2a applies, EU, RDU and CP make incorrect predictions, while CCP makes the correct predictions.

9.2. Composite rank dependent utility theory (CRDU).

It is also possible to explain the stylized facts of insurance by modifying rank dependent utility (RDU).

**Definition 9** (CRDU; al-Nowaihi and Dhami, 2010): If one replaces cumulative prospect theory (CP) with rank dependent utility (RDU) in Definition 8, one obtains composite rank dependent utility (CRDU).

**Proposition 10**: Under composite rank dependent utility theory (CRDU), for a given expected loss, a fraction \( \mu \) of decision makers will not insure, for all sufficiently small probabilities. However, the remaining fraction, \( 1 - \mu \), uses RDU and, so, Proposition 4(a),(b) apply to them. These individuals will insure fully against all losses of sufficiently small probability.

CRDU and CCP would seem to be equally effective for the set of insurance puzzles that we address. This need no longer be the case in other problems. \(^{35}\) Unlike CRDU, CCP (like CP) incorporates reference dependence, loss aversion and richer attitudes towards risk. Hence, CCP explains everything that CRDU can, but the converse is false.

10. Conclusion

A satisfactory theory of insurance should explain why some decision makers buy insurance for low probability losses while others do not. Whether premiums are actuarially fair, unfair or subsidized, and for a fixed expected loss, evidence shows that there is a probability below which the take-up of insurance drops dramatically. In other contexts, decision makers oversubscribe to insurance even when it is not good value. Furthermore, a satisfactory theory should incorporate the finding that ‘context dependence’ and ‘framing’ effects the demand for insurance. Despite commendable progress, existing theories of insurance are unable to account for all of these findings.

EU fails to account for these findings, 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), the two leading alternatives to EU under risk also fail to explain these findings. Indeed, they predict even more insurance than EU when the predicted insurance under EU is already excessive.

We use al-Nowaihi and Dhami’s (2010) composite prospect theory (CCP) to address the insurance puzzles. CCP is axiomatically founded, is consistent with the evidence and explains a wide variety of behavior in diverse contexts. Under CCP, a fraction $\mu$ of decision makers place very low salience on very low probability events. Their behavior fits in with the low take-up of insurance for low probability losses. The remaining fraction $1 - \mu$, behaves as in standard CP models. Standard CP models share with standard RDU models, the feature that low probabilities are infinitely overweighted. This makes the fraction $1 - \mu$ eager to buy insurance against low probability losses.

Furthermore, in CCP, $\mu$ is a parameter of the model. CCP is consistent with the notion that $\mu$ is likely to depend on the context, framing of the problem, emotions etc.

The standard asymmetric information model of insurance predicts that the low risk types could drop out of the market (market failure). In that case, the actual claim rate is higher than the average risks in the population (hypothesis H1 in subsection 2.2). This prediction has been successfully empirically tested. It, however, is also a prediction of the insurance model under CCP. Suppose there is a probability distribution over losses across the population. A fraction $\mu$ who face losses with very low probability are not interested in buying insurance, hence, they drop out leaving a pool of potential insurance buyers who face a higher probability of a loss.

We argue that of the alternatives, we offer the most complete account of the demand for insurance, which is critical in almost all areas of economics.
11. Appendix: Proofs

Proof of Proposition 1: These properties follow 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 \( \lim_{p \to 0} \frac{w(p)}{p} = \lim_{p \to 0} w'(p) \).
(b) If \( w(p) \) is differentiable in a neighborhood of \( p = 1 \), then \( \lim_{p \to 1-} \frac{1-w(p)}{1-p} = \lim_{p \to 1-} w'(p) \).

Proof of Lemma 1: (a) Let \( p \to 0 \). Since \( w \) is continuous (Proposition 1c), \( w(p) \to w(0) = 0 \) (Proposition 1a). By L’Hospital’s rule, \( \frac{w(p)}{p} \to \frac{dw(p)/dp}{dp} = w'(p) \).
(b) Similarly, if \( p \to 1 \), then \( w(p) \to w(1) = 1 \). By L’Hospital’s rule, \( \frac{1-w(p)}{1-p} \to \frac{d(1-w(p))/dp}{d(1-p)/dp} = w'(p) \). ■

Proof of Proposition 2: Straightforward from Definition 4. ■

Proof of Proposition 3: From (5.1) 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 \( \lim_{p \to 0} \frac{w(p)}{p} = \infty \) and, hence, \( \lim_{p \to 0} \frac{w(p)}{p} = \infty \). This establishes (ai). On the other hand, if \( \alpha > 1 \), then \( \lim_{p \to 0} \frac{w(p)}{p} = -\infty \) and, hence, \( \lim_{p \to 0} \frac{w(p)}{p} = 0 \). This establishes (bi). From (5.1) we get \( w'(p) = \frac{\alpha \beta}{p} (-\ln p)^{\alpha-1} w(p) \). If \( \alpha < 1 \), then \( \lim_{p \to 0} w'(p) = \infty \). Part (a) then follows from Lemma 1b. If \( \alpha > 1 \), then \( \lim_{p \to 1} w'(p) = 0 \). Part (b) then follows from Lemma 1b. ■

Proof of Proposition 4: (a) Consider an expected loss
\[ \bar{L} = pL. \] (11.1)
Differentiate (6.1) with respect to \( C \) to get
\[ U'_1(C) = -rw(1-p)u'(W-rC-f) \]
\[ + (1-r)(1-w(1-p))u'(W+(1-r)C-f-L). \] (11.2)

Since \( u \) is (strictly) concave, \( u' > 0 \) and \( 0 < r < 1 \), it follows, from (11.2) that \( U'_1(C) \) is a decreasing function of \( C \). Hence,
\[ U'_1(L) \leq U'_1(C) \leq U'_1(0) \text{ for all } C \in [0, L]. \] (11.3)
Using (3.1), replace \( r \) by \((1+\theta)p \) in (11.2), then divide both sides by \( p \), to get
\[ \frac{U'_1(C)}{p} = -(1+\theta)w(1-p)u'(W-(1+\theta)pC-f) \]
\[ + (1-(1+\theta)p) \frac{1-w(1-p)}{p} u'(W-(1+\theta)pC-f-L+C). \] (11.4)
For $C = 0$ and $C = L$, (11.4) gives (using (11.1)):

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

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

(11.5)

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

$$\frac{U_I'(0)}{p} < \frac{1 - w(1 - p)}{p} u'_{\text{max}} - (1 + \theta) w(1 - p) u' (W - f).$$

(11.6)

From (11.6) and (6.6) we get,

$$\frac{U_I'(L)}{p} = F(p) u' \left( W - (1 + \theta) L - f \right).$$

(11.8)

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

$$U_I'(L) > 0 \Leftrightarrow F(p) > 0.$$

(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

$$f < LF(p) \Rightarrow U_{NI} < U_I(L).$$

(11.10)

Thus, if fixed costs are bounded above by $LF(p)$, the participation constraint is guaranteed to hold. Let

$$q = 1 - p.$$

(11.11)

(b) Suppose $w(p)$ infinitely-underweights near-one probabilities. Then, from (11.11) and Definition 2b, $\lim_{p \to 0} \frac{1-w(1-p)}{p} = \lim_{q \to 1} \frac{1-w(q)}{1-q} = \infty$. Hence, from (6.6), for given expected loss, $L$, we can find a $p_1 \in (0, 1)$ such that, for all $p \in (0, p_1)$, we get $f < LF(p)$. From (11.10) it follows that the participation constraint (6.5) is satisfied for all $p \in (0, p_1)$. From $f < LF(p)$ 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, $\lim_{p \to 0} \frac{1-w(1-p)}{p} = \lim_{q \to 1} \frac{1-w(q)}{1-q} = 0$. 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 5: (a) Since \( v \) is strictly concave, \(-v\) is strictly convex. Hence, from (7.4), 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 (5.1) and the value function (7.1). Consider an expected loss

\[ \bar{L} = pL \]  

From (5.1), (7.1), (7.6), (7.7), (7.9) and (11.12), simple algebra leads to

\[ f < \bar{L}F(p) \Rightarrow V_{NI} < V_I(L). \]  

From (7.9) and Proposition 3a(i), \( \lim_{p \to 0} F(p) = \infty \). 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 (7.6) and (7.7) we get the following

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

\[ \lim_{p \to 0} \frac{V_I(L) - V_{NI}}{p} = v(L) \lim_{p \to 0} \frac{w(p)}{p} - v((1 + \theta)\bar{L} + f) \lim_{p \to 0} \frac{1}{p}. \]

Suppose \( w(p) \) zero-underweights infinitesimal probabilities. Then, from Definition 3a, \( \lim_{p \to 0} \frac{w(p)}{p} = 0 \). 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 6: Straightforward from Definitions 1 and 7.

Proof of Proposition 7: Follows by direct calculation from (8.4) and (8.5).

Proof of Proposition 8: Part (a) follows from part (bi) of Proposition 3, since \( \alpha_0 > 1 \). Part (b) follows from part (bii) of Proposition 3, since \( \alpha_1 > 1 \).

Proof of Proposition 9: The fraction \( 1 - \mu \) of decision makers use CP, hence, their behavior is given by Proposition 5(a),(b). The behavior of the remaining fraction, \( \mu \), follows from Propositions 5c and 8a.

Proof of Proposition 10: For the fraction \( 1 - \mu \), of decision makers Proposition 4(a),(b) applies. The behavior of the remaining fraction \( \mu \), follows from Propositions 4c and 8b.

References


