ESSAYS ON DECISION-MAKING UNDER UNCERTAINTY

DI WANG, MSc.

THESIS SUBMITTED TO THE UNIVERSITY OF NOTTINGHAM

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

SEPTEMBER 2017

Abstract

This thesis consists of three closely related studies investigating individual decision-making under risk and uncertainty, with a focus on decision weighting. Chapter 1 provides an overview of the common themes and theoretical framework for this research.

Chapter 2 reports the development of a simple method to measure the probability weighting function of Prospect Theory (Kahneman & Tversky, 1979) and rank-dependent utility theories. Our method, called the Neo-Lite method, is based on Abdellaoui et al. (2011)'s source method and the Neo-additive weighting function (Chateauneuf et al., 2007). It can be used for both risk (known probabilities) and for ambiguity (unknown probabilities). The novelty of our method lies in how data of decision weights are used to obtain the measurement of the whole function. Compared to the more widely used parametric fitting, our method is simpler, as it minimizes the number of decision weights required and does not rely on the elicitation of subjective probabilities (for ambiguity). An experiment of choice under risk demonstrates the simplicity and tractability of our method. The predictive performance of probability weighting functions measured using our method is shown to be almost equally good to that measured using the standard parametric fitting method.

Chapter 3 presents a theory of choice under risk primarily to explain why individual probability weighting functions are often found to be non-linear and inverse-S shaped. Our rationale for non-linear probability weighting is based on a psychologically grounded feature of choice making, a feature we call *attention-based state weighting*. We show that, under well-defined circumstances, our theory can be equivalent to Cumulative Prospect Theory (Tversky & Kahneman, 1992) with a probability weighting function depending on not only ranks but also sizes of outcomes of risky prospects. This allows our theory to accommodate evidence about probability weighting that cannot be explained by Prospect Theory or Cumulative

Prospect Theory.

The evidence just mentioned refers to recent findings that people have stake-sensitive probability weighting functions. Chapter 4 reports an experiment that further explores the idea of stake-sensitive decision weighting. In addition, the experiment also tests hypotheses derived from the theory presented Chapter 3. In this chapter, we use a more general concept, decision weights, to refer to probability weights that can be stake-sensitive. Particularly we investigate whether and how decision weights are affected by two main properties of a lottery: outcome level (or expected payoff level) and outcome spacing (or the ratio of the best outcome of the lottery to its worst outcome). We elicit subjects' Certainty-Equivalents for carefully-designed sets of lotteries and estimate their decision weights and utility curvatures using a model slightly more general than Cumulative Prospect Theory. Our main finding is that only outcome spacing has significant and systematic influence on decision weighting at the aggregate level, and that both outcome level and outcome spacing have systematic and significant effects at individual levels. This finding, together with the theory of Chapter 3, challenges the common understanding of probability weighting.

Chapter 5 concludes the thesis by summarizing all findings in previous chapters, discussing the implications, and pointing to directions for future research.

Acknowledgements

I am deeply grateful to my supervisors Robin Cubitt and Chris Starmer. It has been an honor to be their student. They have taught me consciously and unconsciously how good research is done. I appreciate all their contributions of time and ideas to the development of the thesis, especially during tough times of this pursuit.

I would like to thank Centre for Decision Research and Experimental Economics (CeDex), for contributing immensely to my personal and professional development. The research centre has been a source of friendships as well as good advices and collaborations. I am particularly indebted to Simon Gächter, Daniele Nosenzo, Martin Sefton, Dennie van Dolder, and Silvia Sonderegger. I also owe my thanks to Jose Guinot Saporta, Despoina Alempaki, Lingbo Huang, Xueheng Li, and Cindy Fu, who have helped with my experiments in various ways. I would like to thank my colleagues Zhixian Yu, Lu Dong, Till Weber, Hanna Fromell, Vasileios Kotsidis, Orestis Kopsacheilis, Valeria Burdea, Ruslan Shichman, and many others, for making the PhD years a pleasurable experience. Many stimulating conversations with people I met in conferences and other various occasions have also contributed to my thesis. They include (but not limited to) Peter Wakker, Robert Sugden, Mohammed Abdellaoui, Uzi Segal, and Yoram Halevy.

I gratefully acknowledge Nottingham University School of Economics for partly funding my study, and CeDEx as well as NIBS for funding my experiments. I owe many thanks to Sarah Nolan and Gianni De Fraja, who have kindly provided support for me when I need it most.

Lastly, I would like to thank my family for all their love and encouragement, particularly my parents who have been supporting me in all my pursuits.

<div style="text-align: right">

Di Wang

September 2017

</div>

Contents

List of Tables

List of Figures

Chapter 1: Introduction

Theories of individual decision-making play a fundamental role in economic theory. Most economically relevant decisions have to be made in the presence of uncertainty. Since Knight (1921), situations of uncertainty are normally thought to contain two different scenarios: decision under risk where probabilities of uncertain events are objectively given and known, and decision under ambiguity where the probabilities are unknown[1]. This thesis is a collection of three studies (Chapters 2, 3 and 4) reporting research that contributes mainly to the area of decision-making under risk. While each of the three chapters is self-contained and can be read independently from the others, there is a common topic for all three studies: *probability weighting and decision weighting*[2].

The two concepts become widely known since the seminal paper of Prospect Theory (PT hereafter, Kahneman & Tversky, 1979), which models how people make choices between lotteries with known probabilities. PT has been one of the most important theories of decision making under risk in the past decades and has been applied in a wide variety of contexts (Wakker 2010). It generalizes the classical Expected Utility Theory (EUT hereafter, Von Neumann & Morgenstern, 1945) by introducing probability weighting, reference-dependent evaluation of outcomes, and loss aversion, and provides systematic explanations for the major deviations from EUT such as the Allais paradox (Allais, 1953), the certainty effect, and framing effects (Kahneman and Tversky, 1979, 1985).

In PT, lotteries are modeled as risky prospects which are often denoted as $(p_1:x_1, \ldots, p_n:x_n)$, that is, yielding outcome $x_i \in \mathbb{R}$ with probability p_i, where $i =1$,

---

[1] Knight (1921) used the term 'uncertainty' to describe the latter case. It is in later literature what Knight called uncertainty is commonly labeled as ambiguity.

[2] Since the subtle difference between these two concepts is important for this thesis, we will explain this difference carefully later in this chapter.

2, …, n. Preferences over the prospects are then represented by the function

$$V(p_1: x_1, \dots, p_n: x_n) = \sum_{i=1}^{n} w(p_i) v(x_i)$$

where $w: [0,1] \to [0,1]$ is called the *probability weighting* function which transforms objective probabilities into probability weights, and $v: \mathbb{R} \to \mathbb{R}$ is a utility (or value) function that measures the decision maker's subjective value of outcomes[3]. Empirical studies of the probability weighting function have been numerous (e.g. Tversky & Kahneman 1992, Wu & Gonzalez 1996, Bleichrodt & Pinto 2000) and have typically found it to be increasing, non-linear and inverse-S shaped, that is, $w(p) > p$ for small probabilities and $w(p) < p$ for moderate or larger probabilities. This non-linear probability weighting is often interpreted as reflecting an intrinsic psychological element of people's diminishing sensitivity towards less-extreme probabilities (e.g. Kahneman & Tversky 1979, Tversky & Wakker 1995, Prelec 1998).

Since EUT, the value of a lottery can be seen as a weighted average of the values of its outcomes. The weights attached to values of outcomes are often called *decision weights*. In EUT, decision weights are objective probabilities, whereas in PT, decision weights are probability weights, which can be different from objective probabilities. The latter implies that, for PT, decision weights are *solely* determined by the probability weighting function, which captures a decision maker's misperceptions of likelihoods or her intrinsic attitude towards probabilities. As recognized by Tversky & Kahneman (1992), this may be too restrictive to capture the effect of important contextual factors on decision weighting. In addition, another feature of the PT model that has been a subject of criticism is that it allows violations of 'stochastic dominance' (Kahneman & Tversky 1979, pp. 283-284), which requires that a shift of probability mass from bad outcomes to better outcomes leads to an

---

[3] In PT, an outcome can be either a gain or a loss based on the decision maker's reference point, and the value function $v(.)$ can be different for gains and for losses. Although PT does not restrict outcomes to be money, this thesis considers only monetary outcomes.

improved prospect (Fennema & Wakker 1997).

These theoretical problems have been dealt with in a new version of PT, called Cumulative Prospect Theory (CPT henceforth, Tversky and Kahneman 1992)[4]. In particular, CPT satisfies stochastic dominance and incorporates the rank-dependence feature of decision weighting that was introduced by Quiggin (1982) and Schmeidler (1989)[5]. As a result, decision weights in CPT depend not only on probability weights, but also on ranks of the corresponding outcomes among all possible outcomes of the lottery[6]. For the outcome domain of gains, the preference function of CPT can be written as

$$V(p_1:x_1, \ldots, p_n:x_n) = \sum_{i=1}^{n} d_i v(x_i)$$

where $x_1 > \cdots > x_n$ and the decision weights d are defined by:

$d_i = w(p_i)$, for $i = 1$

$d_i = w(p_1 + \cdots + p_i) - w(p_1 + \cdots + p_{i-1})$, for $2 \leq i \leq n$

where w is the probability weighting function as in PT that captures people's intrinsic attitude towards probabilities. CPT generalizes EUT and PT, because EUT is a special case of CPT when $d_i = p_i$ and the characterization of PT that is captured by the formulation above becomes a special case of CPT when $d_i = w(p_i)$. Hence, decision weight is a more general concept than probability weight, in the sense that the former refers to whatever is attached to the value of lottery outcomes and can be context-dependent[7], whereas the latter is often thought as the part of decision weighting that reflects context-independent attitude towards objective probabilities

---

[4] Models of similar forms were introduced by Starmer & Sugden (1989) and Luce & Fishburn (1991).

[5] Another improvement is that CPT also extends to the analysis of both decision under risk and decision under ambiguity. For a survey of non-expected utility, see Slovic et al. (1988), Camerer (1992), and Starmer (2000).

[6] The plausibility of the cumulative form of rank-dependence and the intuition of rank-dependent decision weights have been explicitly discussed in Wakker (1989) and Diecidue & Wakker (2001) respectively.

[7] Context-dependence here, and for the rest of the thesis, refers to the dependence on all sorts of elements of the decision context, such as lottery outcomes, decision tasks, source of uncertainty, etc., *except objective probabilities*. Context-dependent decision weights are weights that does not only depend on objective probabilities of the outcomes, but also on contextual elements.

due to psychological reasons such as human cognitive limitations[8].

While allowing decision weights to depend on the rank of lottery outcomes may increase the descriptive power of CPT, it certainly complicates the empirical measurement of the model (Abdellaoui 2000). However, since the way decision weights depend on probability weights are pre-specified in CPT, the only component of decision weights to be measured is the probability weighting function, which is thought to be a stable personal component of individual decision making under risk. And given that, in CPT, the decision weight of the best outcome of a lottery is still the same as its probability weight, the probability weighting function could be mapped with measurements of the decision weights for the *best* lottery outcome as the probability of that outcome varies from 0 to 1. This lays the theoretical ground for measuring the probability weighting function through the elicitation of the decision weights of the best lottery outcomes (e.g. Wu & Gonzalez 1996, Abdellaoui 2000, Abdellaoui et al. 2011).

The first study of this thesis has two goals. First we present a method to measure the probability weighting function. As CPT gains its prominence among decision theories, empirical studies, including those measuring CPT functions, becomes increasingly important. For example, the application of CPT to many aspects of the real world needs to be based on proper measurements of the probability weighting function (e.g. Bleichrodt et al. 1999, Bleichrodt & Pinto 2000). Since CPT provides a model for both the case of choice under risk and the case of ambiguity, it would be desirable to have a simple and tractable method which can measure the probability weighting function in both cases[9]. The method we introduce is called the Neo-Lite method. Unlike most existing methods, our method does not require data fitting techniques. Nor does it require the elicitation of subjective probabilities for ambiguity, which can be tricky because often assumptions about how probabilities are assigned to ambiguous events are needed (see e.g. Ellsberg,

---

[8] See e.g., Edwards (1954, 1962), Tversky & Kahneman (1983), Brandstätter et al. (2002), Trepel et al. (2005).
[9] Many methods from the literature achieve this, e.g. Wakker & Deneffe (1996), Gonzalez & Wu (1999), Abdellaoui (2000), Van de Kuilen & Wakker (2011). Chapter 2 provides a more detailed review of existing methods.

1961).

The second purpose of the first study is to test experimentally the predictive power of probability weighting functions measured with different methods and with different functional forms. Specifically, we want to compare our Neo-Lite method with the standard parametric fitting method, and to compare the non-linear class of weighting functions (e.g. Goldstein & Einhorn 1987, Prelec 1998) with the linear class (Chateauneuf et al. 2007). We elicited individuals' Certainty-Equivalents (CE henceforth) of a set of lotteries as well as their choices between pairs of lotteries. Using the CEs we measured probability weighting functions using the two methods and various functional forms for each individual. The probability weighting functions measured were then used to make predictions about the choices. Our results show that the Neo-Lite method works as a good substitute for parametric fitting with the Neo-additive weighting function (Chateauneuf et al. 2007), but has its limitations as a substitute for parametric fitting with a nonlinear probability weighting function. In terms of predictive performance, we found that the non-linear probability weighting functions proposed by Goldstein & Einhorn (1987) and Prelec's (1998) slightly outperform the Neo-additive function.

Another important finding of the first study is that the probability weighting function measured using CE data is dramatically different from that measured using binary choice data. Although this is perhaps unsurprising in light of the well-known phenomenon of 'preference reversal' (e.g. Lichtenstein & Slovic 1971, Grether & Plott 1979, Seidl 2002, Cubitt et al. 2004), it reinforces the idea that, for theories of choice under risk that employ a decision weighting function and a value function, allowing decision weights to depend on non-probabilistic factors[10] can considerably improve the descriptive power of the theories. This idea deeply motivated the second study of this thesis.

Our second and third studies start to question the CPT model, particularly its modelling of decision weights. We are also motivated by the series of findings that

---

[10] Here probabilistic factors basically refer to probabilities and probability weights, and in the case of preference reversal, the key non-probabilistic factor is the type of task used to elicit preferences.

decision weights are affected by lottery outcomes, particularly whether outcomes are perceived as gains or losses (e.g. Tversky & Kahneman 1992, Abdellaoui et al. 2005) and the level of outcomes (Etchart-Vincent 2004, 2009; Fehr-Duda et al. 2010). In our second study (Chapter 3), a theory of decision-making under risk is presented, primarily to account for the dependence of decision weights on lottery outcomes and the preference reversal phenomenon (or the dependence of decision weights on the type of tasks used to elicit preferences). It is natural to think of a more complicated decision weight function for CPT. However, we chose an alternative approach. Our theory is based on the concept of the Savage-act (Savage, 1972), the framework of which is fundamentally different from that of PT or CPT. An important feature of the latter two is that lotteries, treated as prospects, are evaluated *independently* and the evaluations determine the results of choice-making. In contrast, Savage's framework provides a natural way to build in important factors such as cross-act comparisons, that is, allowing the value of a lottery to depend on the alternatives to choose between (e.g. Loomes & Sugden 1982, 1986; Bordalo et al. 2012).

The key outputs of the second study are the followings. First, building on a psychological intuition similar to Bordalo et al. (2012), our theory implies CPT with an inverse-S shaped probability weighting function (or the decision weight attached to the best outcome of a lottery as a function of the probability of the best outcome). In other words, we show that the inverse-S shape of the typically found probability weighting function can come from a decision-making mechanism that does not pre-suppose any form of probability weighting. Secondly, the implied probability weighting function also depends on lottery outcomes and can be different for gains and losses. We then show that this can accommodate the empirical findings mentioned earlier regarding the outcome-dependence of decision weights. Thirdly, our theory can imply substantially different probability weighting functions (in terms of their shapes) for money valuation tasks and for binary choice tasks under risk, thus can potentially explain the 'preference reversal' phenomenon. We provide intuitive analysis of this in Chapter 3 and analytical support for this in Appendix 3.4. In general, our theory can do what CPT could do with a probability weighting function

which depends on multiple non-probabilistic factors, but is more parsimonious and intuitive.

In Chapter 4, we explore one of the key implications of our theory by studying experimentally the effect of lottery outcomes on decision weighting. As mentioned by Tversky & Kahneman (1992), both the level of outcomes and the spacing of them can have an effect on decision weights. For two-outcome lotteries, the latter basically refers to the difference between the best and worst outcomes of the lottery. There have not been many studies that investigate explicitly the effect of outcome level or outcome spacing or both on decision weights, though some evidence on the existence of both types of effect can be found in Camerer (1992). More recently, Etchart-Vincent (2004, 2009) studied this question for the domain of losses and found that, compared to small-loss gambles, large-loss gambles are associated with smaller decision weights of the worst outcome. Another study by Fehr-Duda et al. (2010) found that in the domain of gains, decision weights of the best lottery outcome decrease as the general stake level of the lotteries increase, and that in the domain of losses, stake level has little effect on decision weighting. On the one hand, our third study, which focuses on the effects of outcome level *as well as outcome spacing,* fills a gap in the literature by studying explicitly the two effects. On the other hand, compared to Etchart-Vincent (2004, 2009) and Fehr-Duda et al. (2010), we have a stronger experimental control to isolate the effect of outcome level from the effect of outcome spacing. This is important because the two factors interact with each other and both may affect decision weighting. Based on our intuition and on one of the key predictions from our theory, we hypothesize that outcome spacing has a significant effect on decision weighting and plays a more important role in determining decision weights than outcome level does.

We elicited CEs from subjects and fitted data to a generalization of CPT with a focus on the change of decision weights as lottery outcome spacing and outcome level change. Our main results are that, in general, outcome spacing has a systematic and significant effect on decision weighting at both aggregate and individual levels, whereas outcome level only has a significant effect at individual level. For lotteries

with the same probability distribution over outcomes, the estimated decision weights of the best lottery outcome are smaller for lotteries with larger outcome spacing, and also smaller for lotteries with higher outcome levels. The latter is consistent with Fehr-Duda's (2010) results and potentially explains finding that risk aversion increases as stakes are higher (e.g. Binswanger 1981, Holt & Laury 2002). These results also lend support to the theory presented in Chapter 3.

Since PT, the theoretical and empirical research of decision weighting under risk has attracted enormous attention from psychologists and economists. Yet, our knowledge of this topic may still be far from complete, because there could be other non-probabilistic factors to be discovered[11] which systematically influence decision weights. The three studies (Chapter 2, 3, and 4) are mainly devoted to advancing our understanding of probability weighting and decision weighting for risk, and to providing a theoretical development as a better alternative to CPT when the descriptive power of the latter becomes weaker as evidence accumulates. This thesis ends with Chapter 5, which summarizes the key messages from the previous chapters, provides a more specific discussion of the contributions of our three studies, and points to potentially fruitful new directions for future research.

---

[11] For example, decision weighs have also been found to be affected by factors such as individual demographic characteristics (e.g. Fehr-Duda et al. 2006; Harbaugh et al. 2002), on emotional state (e.g. Fehr et al. 2007), on the emotional or affective content of the payoffs (e.g. Rottenstreich and Hsee 2001), and on the sources of uncertainty (e.g. Kilka & Weber 2001, Abdellaoui et al. 2011).

Chapter 2: A Simple Method to Measure Probability Weighting Functions and its Experimental Implementation

2.1 Introduction

It has been understood, since Allais (1953), that Expected Utility Theory (EUT, von Neumann & Morgenstern 1945) has its limitation in describing individual behaviors under risk. A huge body of literature has documented violations of EUT (see e.g. Starmer, 2000 for a survey). Among non-EU theories, Prospect Theory (PT, Kahneman & Tversky 1979) is one of the most important (Wakker, 2010). In PT, most of the violations of EUT are accommodated by a (non-linear) probability weighting function, which is often regarded as one of the most important theoretical developments of behavioral economics.

In terms of empirical research, an extension of PT, Cumulative Prospect Theory (CPT, Tversky and Kahneman 1992), has gained its importance and popularity. It generalizes the case of decision under risk and decision under ambiguity[1], and introduces the key feature of rank-dependent utility theories (e.g. Quiggin 1982, Schmeidler 1989) that decision weights do not only depend on probability weights but also the rank of lottery outcomes. Nevertheless the probability weighting function in CPT still plays a crucial role in explaining deviations from EUT, capturing individual risk attitudes (when the probabilities of the events are known),

---

[1] Decision under risk relates to situations where known statistical probabilities are available for uncertain events. It has been understood since Keynes (1921) and Knight (1921) that statistical probabilities are often unknown or unavailable. For such cases, referred to as ambiguity, Savage (1954) proposed Subjective Expected Utility Theory (SEUT), where prospects are evaluated by their (subjective) probability weighted average utility. Ellsberg (1961) put forward a more fundamental problem: it seems that often people cannot assign probabilities, not even subjective ones, to random events (ambiguity). Ellsberg's paradox puts the preference basis of SEUT into question.

subjective beliefs and ambiguity attitudes (when probabilities of the events are unknown). Developing tools to measure this function hence becomes important for both empirical research on decision under uncertainty and the application of the probability weighting function to real decision analysis.

Various methods have been developed (e.g. Wakker & Deneffe, 1996; Gonzalez & Wu, 1999; Abdellaoui, 2000; van de Kuilen & Wakker, 2011) that can measure parametric-free decision weights under both risk and ambiguity. It is sometimes also useful to obtain measurements of parameterized probability weighting functions for purposes of both theoretical research and empirical study (such as predictive tests). This could be easy for the case of risk but tricky for ambiguity, because the latter normally requires the elicitation of subjective probabilities that enable the mapping of ambiguous events to probability weights. In many studies, measurement of the probability weighting function for ambiguity relied on stated subjective probabilities (e.g. Heath & Tversky, 1991; Fox & Tversky, 1998). Recently, the revealed-subjective-probability approach is more often used (e.g. Abdellaoui et al. 2005, 2011) to avoid potential biases from stated subjective probabilities.

This chapter first introduces a simple method, called the Neo-Lite method, to measure probability weighting functions of CPT, for both risk and ambiguity. More precisely, Neo-Lite provides a convenient way to use decision weights to obtain the probability weighting function. This method can be more efficient than previous measurement methods because it minimizes the amount of data (decision weights) required and greatly simplifies the processing of data. Moreover, unlike most existing methods, our method does not require the elicitation of subjective probabilities for ambiguity in the sense that the parameters of the probability weighting function can be obtained without any inference about the underlying subjective probabilities of the uncertain events.

Since our method is based on the assumption of the Neo-additive probability weighting function (Chateauneuf et al., 2007), a second purpose of this study is to experimentally compare the predictive power of different classes of probability weighting functions widely used in the literature, particularly the non-linear class

represented by Goldstein and Einhorn (1987) and Prelec 1998), and the Neo-additive (or linear) class (Chateauneuf et al., 2007). For the latter, we apply both the standard parametric fitting method and the Neo-Lite method to measure the probability weighting function and to see how the two measurements predict.

In general, we find that the Neo-Lite method works as a good substitute for parametric fitting with the Neo-additive weighting function, both descriptively and predictively. Conditional on using a linear function, our method succeeds in revealing a function which is essentially indistinguishable from that generated with parametric fitting method. To compare the predictive power of the linear and non-linear classes of probability weighting functions measured from our experiment, we conduct both out-of-sample and in-sample predictive tests. We find that the performances of the two classes of weighting functions are very similar in the out-of-sample test, but the non-linear class outperforms the linear one in the in-sample test. This suggests that the latter class of weighting function has its limitation in describing and predicting choices under risk.

This chapter proceeds as follows. Section 2.2 briefly introduces CPT, the theory under which our method works. Section 2.3 provides a brief review of the existing measurement methods in the literature. Section 2.4 presents the Neo-Lite method. Experimental design and results are presented in Section 2.5. Section 2.6 concludes.

2.2 Prospect Theory for Risk and Ambiguity

In this section we introduce basic concepts, notations and the theoretical framework based on Tversky & Kahneman (1992) and Abdellaoui et al. (2011). For simplicity, we use a similar set of concepts and notations as used by Abdellaoui et al. (2011). Let $X \in \mathbb{R}$ be the set of outcomes and S be the state space or the set of mutually exclusive uncertainty events E. Let $(E_1: x_1, \dots, E_n: x_n)$ denote a prospect yielding outcome x_j if event E_j happens, with $j =$1, 2, …n. For the purpose of this chapter,

outcomes are restricted to be gains of money, although this is not necessary for the new method we propose. A prospect is uncertain when only one of the events E_1, $E_2, ..., E_n$ will happen but it is unknown ex ante which one will happen. Since in our measurements we only need two-outcome positive prospects $(E:x,$ not $E:y)$, we use the simpler notation $x_E y$ for such prospects, and by this notation we assume that $x > y \geq 0$. The Certainty-Equivalent (CE) of a prospect is the sure amount that makes the decision maker indifferent to the prospect.

A decision theory for uncertainty specifies how the decision maker's preferences over prospects are determined by a value function V. For instance, Expected Utility holds if a prospect $(E_1:x_1, ..., E_n:x_n)$ is evaluated by $V = \sum_{j=1}^{n} p(E_j) v(x_j)$, with v, the utility function (also called value function) of outcomes, continuous and strictly increasing and $p(E_j)$ the probability of event E_j. As in many other empirical studies on weighting functions, we use CPT for two-outcome prospects, which coincides with Quiggin's (1982) rank-dependent utility because we only consider positive outcomes. It holds if there exists a utility function $v: \mathbb{R} \to \mathbb{R}$ and an event weighting function W such that preferences maximize

$$V(x_E y) = W(E)v(x) + \big(1 - W(E)\big)v(y) \tag{2.1}$$

where W assigns a number $W(E)$ between 0 and 1 to each event E, such that $W(\emptyset) = 0$, W is 1 at the universal event, and $E \supset F$ implies $W(E) \geq W(F)$. If the probabilities of the events are known, then the prospects are evaluated by

$$V(x_E y) = w(p(E))v(x) + \big(1 - w(p(E))\big)v(y) \tag{2.2}$$

where $w(.)$ is a probability weighting function[2] that maps $[0,1]$ to $[0,1]$ and is strictly increasing and continuous, with $w(0) = 0$ and $w(1) = 1$.

When probabilities are unknown or unknowable (i.e. under ambiguity), we are

---

[2] For clarification, we use lower case $w(.)$ to denote probability weighting functions for known probabilities exclusively.

back to equation (2.1). This generates great empirical difficulty in defining and measuring the weighting function for ambiguity. Therefore, the development of the source method by Abdellaoui et al. (2011) is important because it provides us a tool to measure the weighting function W defined on subjective probabilities. They have formalized the definition of source of uncertainty[3] and have shown, complementary to Ellsberg (1961)'s finding, that whenever the source of uncertainty is uniform to the decision maker (i.e. all the events from that source have a uniform degree of ambiguity), the decision maker can be seen as having a (subjective) probability measure p over the event space and a function w_Q, carrying subjective probabilities to decision weights, such that for any event E from source Q we have

$$W(E) = w_Q(p(E)) \tag{2.3}$$

where Q stands for the set of sources of uncertainty[4]. This generalizes Prospect Theory for risk because risk can be a special source of uncertainty. Now equation (2.1) becomes

$$V(x_E y) = w_Q(p(E))v(x) + \left(1 - w_Q(p(E))\right)v(y) \tag{2.4}$$

In what follows we consider only two-outcome lotteries and assume that individual preferences for uncertain binary prospects are represented by equation (2.4) unless otherwise specified.

2.3 Existing Methods in the Literature

There have been generally three classes of empirical methods to measure the probability weighting function for risk: parametric fitting, non-parametric fitting, and

---

[3] According to Abdellaoui et al. (2011), a source of uncertainty concerns a group of events that is generated by a common mechanism of uncertainty. In their paper, sources are modeled as algebras containing the universal event (certain to happen), the vacuous event (certain not to happen), the complement of each of their elements, and the union of each pair of their elements. In this chapter, we also take sources as algebras.
[4] We use Q instead of S to denote the set of uncertainty sources because the latter is used to represent 'State'.

non-parametric non-fitting[5].

Parametric Fitting

Needless to say, parametric fitting has been a standard method being widely used in empirical economics, due to the development of econometrics. A substantial part of empirical studies on choice under risk have used parametric fitting to measure the probability weighting function for various research purposes (e.g. Camerer and Ho 1994, Wu & Gonzalez 1996, Bleichrodt and Pinto 2000, Kilka & Weber 2001, Blavatskyy 2005, Booij et al. 2010).

For more specific examples, Tversky and Kahneman (1992) found direct evidence supporting the presence of probability weighting in decision making under risk, using parametric fitting to measure the probability weighting function. Hey and Orme (1994) used parametric fitting to test different theories of choice under risk and found that theories of the rank-dependence family perform relatively better than other families. More recently, Fehr-Duda et al. (2010) studied how risk tolerance varies with stake sizes. They fitted the value function and probability weighting function of Prospect Theory (Tversky & Kahneman, 1992) and found that, for gains, risk aversion grows as stakes increase. A sub-class of parametric fitting, called the semi-parametric fitting, works by assuming a parametric form of only the objective function, and estimating the parameters of the objective function and the point values of other functions. This method was first proposed by Abdellaoui et al. (2008), and has been gaining popularity due to its efficiency (compared to standard parametric fitting).

Non-Parametric Fitting

Another measurement technique for probability weighting and utility functions can

---

[5] The book by Peter Wakker (2010) can be a good source for introductions to many of the widely used empirical methods for weighting function measurement.

be called non-parametric fitting (e.g. Gonzalez & Wu, 1999; Abdellaoui et al., 2007, 2011; Stott, 2006). It uses data fitting but does not commit to any parametric family. Instead, it takes every value of the objective function as a parameter. For example, Gonzalez & Wu (1999) proposed a non-parametric method for estimating the probability weighting function for risk and confirmed the inverse-S shaped weighting function with their experimental data. A similar method was also used by Stott (2006) to discriminate, using experimental data, among various functional forms of the stochastic Prospect Theory models. Although not committing to any parametric functional forms can be an advantage for some research purposes, this method normally requires a large amount of data. Abdellaoui et al. (2008) and Van de Kuilen & Wakker (2011) provide a more detailed introduction and discussion of non-parametric fitting.

Non-Parametric Non-Fitting

Methods that use neither parametric functional forms nor data fitting techniques belong to this class. Most of these methods use specifically designed choice questions to elicit points of interests of the objective function. For example, Abdellaoui (2000) developed a method that has two steps: the first step consists of constructing a sequence of choice questions that elicit a sequence of outcomes equally spaced in utility by means of the trade-off method initially proposed by Wakker and Deneffe (1996); the second step uses the sequence of outcomes to obtain a sequence of probabilities equally spaced in terms of probability weighting. Alternative methods that use similar techniques include, e.g., Abdellaoui and Wakker (2005), Blavatskyy (2006), Abdellaoui et al. (2007), and van de Kuilen and Wakker (2011). These methods can generally be used to measure probability weighting midpoints which can reveal the shape of the weighting function. For example, Wu et al. (2005) applied Abdellaoui (2000)'s version of the trade-off method to discriminate between the original version of the PT (Kahneman & Tversky, 1979) and the CPT (Tversky & Kahneman, 1992) and found that the PT performs better.

For empirical studies on choice under ambiguity, similar methods of these classes have been used to measure the (subjective) probability weighting function, but often in more complex ways due to the need for the elicitation of subjective probabilities. For example, Hey et al. (2010) and Kothiyal et al. (2014) both tested the predictive performance of different models of choice under uncertainty and used parametric fitting with the subjective probabilities estimated as extra parameters. Another representative study using parametric fitting for choice under ambiguity is the paper by Abdellaoui et al. (2011). They found experimental evidence supporting the "source-dependence" nature of probability weighting functions. To obtain the weighting functions for ambiguity, they carried out complex procedures. They first used the semi-parametric method to measure the utility function, then obtained decision weights, elicited corresponding subjective probabilities from a sequence of choice questions, and finally did parametric fitting based on the measured decision weights and elicited subjective probabilities. Diecidue et al. (2007) tested the existence and nature of rank dependence for Rank-Dependent theories of choice under uncertainty, employing a non-parametric non-fitting method to elicit the decision weights of ambiguous events without eliciting subjective probabilities. Their method can be a good complement for our method in measuring the weighting function because their method can, as they have shown in their study, elicit decision weights in a simple way, and our method provides a simple way to obtain the parameters of the Neo-additive weighting function from the decision weights.

2.4 The Neo-Lite Method

Our method starts by assuming that the weighting function is Neo-additive (Chateauneuf et al. 2007), or loosely speaking, linear. Although the most common finding about weighting function is the inverse-S shape, the linear one can be seen as an approximation of the inverse-S shape. Indeed, several papers have argued for the

importance of the neo-additive family (e.g. Gilboa 1988; Cohen 1992; Loomes et al. 2002; Teitelbaum 2007; Chateauneuf et al. 2007), and some studies have also found direct evidence of individuals exhibiting weighting functions very close to linear (e.g. Abdellaoui et al. 2011, Figure 4c, Figure 10 & 11). A major advantage of this family is its theoretical parsimony and resulting practical simplicity. Another important advantage is that interpretation of its parameters is clearer and more straightforward than other families (Wakker 2010, p.210). We will show this below.

To show how our method can be applied to decision under ambiguity where the (subjective) probability weighting function can depend on the source of uncertainty, we will work with a source-dependent neo-additive weighting function, i.e.

$$w_Q(p) = \begin{cases} 0, & p = 0 \\ \mu + \gamma p, & 0 < p < 1 \\ 1, & p = 1 \end{cases}$$

where $0 \leq \mu < 1$, $0 \leq \gamma \leq 1 - \mu$, and Q is the source of uncertainty and it can be either risky or ambiguous.
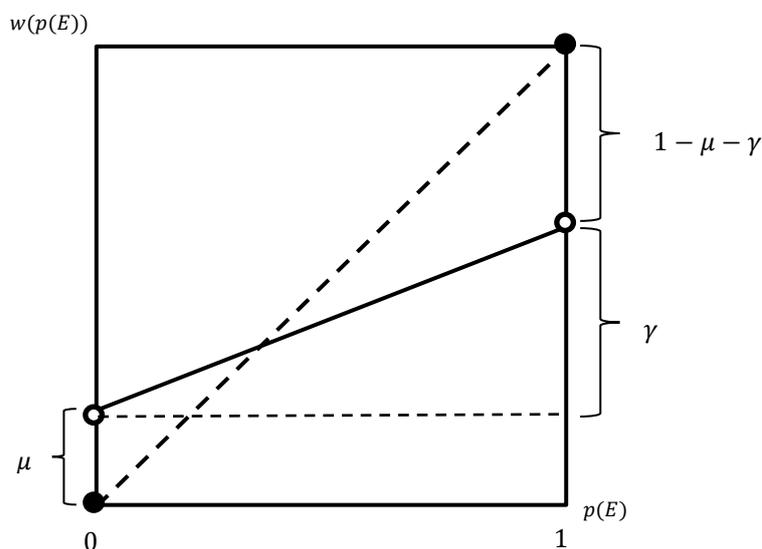
Figure 2.1: A typical Neo-additive probability weighting function (the solid line)

As shown in Figure 2.1, μ corresponds to the lower intercept and γ

corresponds to the slope. Even without looking at the graph it is straightforward to read off the shape of the weighting function from the parameter values. Moreover, two important features of probability weighting, likelihood-insensitivity and pessimism/optimism (Wakker, 2010), can be captured by extremely simple representations of μ and γ. The slope can be seen as an index of likelihood-sensitivity (in other words, $1 - \gamma$ the index of likelihood-insensitivity) because it reflects how much the decision weight changes when the probability changes by one unit. The aggregate pessimism/optimism can be measured by $(1 - \mu - \gamma) - \mu$, i.e., the difference between the upper intercept and the lower intercept, because this index reflects generally to what extent probabilities are underweighted or overweighted (Abdellaoui et al. 2011).

Once the weighting function is restricted to be Neo-additive, it is straightforward to obtain its parameter measurements. For any two disjoint events A and B from S that are non-null and non-universal and the union of which $(A \cup B)$ is non-universal, we have

$$w_Q(p(A)) = \mu + \gamma p(A) \tag{2.5}$$

$$w_Q(p(B)) = \mu + \gamma p(B) \tag{2.6}$$

$$w_Q(p(A \cup B)) = \mu + \gamma p(A \cup B) \tag{2.7}$$

Manipulating equations (2.5), (2.6) and (2.7), we have, since $p(A) + p(B) = p(A \cup B)$,

$$\mu = w_Q(p(A)) + w_Q(p(B)) - w_Q(p(A \cup B)) \tag{2.8}$$

The right-hand side of equation (2.8) is an index of lower-subadditivity (Tversky & Wakker 1995)[6]. This equation shows that, when the weighting function is Neo-additive, the intercept μ is a measure of the degree of lower-subadditivity. In

---

[6] According to Tversky & Wakker (1995), *lower-subadditivity* is a property of a probability weighting function. They defined *lower-subadditivity* with an inequality relation between certain decision weights: a probability weighting function w satisfies lower-subadditivity if $w(p_1) \geq w(p_1 + p_2) - w(p_2)$, for probabilities p_1 and p_2 such that $p_1 + p_2 < 1$.

other words, the right hand side of equation (2.8) captures the boundary effect near zero probability. For example, $w_Q(p(A)) + w_Q(p(B)) > w_Q(p(A \cup B))$ would imply the "possibility effect" that an event has greater impact when it turns impossibility into possibility than when it merely makes a possibility more likely. Given that we have obtained an equation of μ, one way to obtain γ is to measure the degree of upper-subadditivity[7], i.e. to measure the upper intercept. Alternatively, γ can also be obtained if we know the difference between the upper intercept and the lower intercept, which is equivalent to the index of pessimism/optimism. We notice that the index of pessimism/optimism is also easy to measure.

Let $\overline{A}$ be the complementary event of A in source Q. We have

$$w_Q(p(\overline{A})) = \mu + \gamma p(\overline{A}) \qquad (2.9)$$

Adding equation (2.5) and (2.9), with $p(A) + p(\overline{A}) = 1$, we have

$$2\mu + \gamma = w_Q(p(A)) + w_Q(p(\overline{A}))$$

and hence

$$1 - 2\mu - \gamma = 1 - w_Q(p(A)) - w_Q\left(p(\overline{A})\right) \qquad (2.10)$$

which is the index of pessimism/optimism. The right hand side of equation (2.10) measures the degree of pessimism/optimism because it can be re-written as $(p(A) - w_Q(p(A))) + (p(\overline{A}) - w_Q\left(p(\overline{A})\right))$. In other words, if a decision maker overweights any event as much as he underweights the complement of that event (i.e. $p(A) - w_Q(p(A)) = w_Q(p(\overline{A})) - p(\overline{A})$), he can be seen as having no systematic tendency of pessimism or optimism towards probabilities.

Combining equation (2.8) and (2.10), we have

---

[7] See also Tversky & Wakker (1995).

$$\gamma = w_Q\big(p(\overline{A})\big) + 2w_Q\big(p(A \cup B)\big) - w_Q\big(p(A)\big) - 2w_Q\big(p(B)\big) \tag{2.11}$$

Although the equation of γ is less obvious, it is obvious that once we pin down the lower intercept and the difference between the upper intercept and lower intercept (i.e. the aggregate pessimism/optimism), the slope is also pinned down.

Our method is general in the sense that it can be applied to any situation of risk and (uniform) ambiguity where the four uncertain events $A, B, A \cup B,$ and $\overline{A}$ can be constructed. It provides a non-fitting way to measure the parameters of the weighting functions given that we know the four decision weights that appear on the right hand sides of (2.8) and (2.11). One way to obtain the decision weights is to design an experiment that enables either direct elicitation of the decision weights (e.g. Diecidue et al.,2007; van de Kuilen & Wakker, 2011) or elicitation of both utility and decision weights (e.g. Abdellaoui, 2000; Abdellaoui et al., 2008; Kothiyal et al., 2014).

Before measuring the four decision weights, it is important to decide how to specify the four relevant events $A, B, A \cup B,$ and $\overline{A}$. To illustrate this, consider an Ellsberg-type urn containing four colors of balls: red, green, yellow, purple. A ball is to be randomly drawn from the urn to determine the results of some gambles on this urn. One way to specify the four events is the following:

Event *A*= "the ball drawn is red"

Event *B*= "the ball drawn is green"

Event *A*∪*B* = "the ball drawn is either red or green"

Event $\overline{A}$ = "the ball drawn is not red"

There is obviously more than one way of specification, as long as none of the four events end up being null or universal. In the case of risk, where probabilities of all the events are known, Event *A* and Event *B* can have the same probability, reducing the minimum number of decision weights required from four to three.

In practice, however, we don't observe decision weights directly. Decision weights can be obtained by using the methods mentioned in Section 2.3, or by

measuring the utility function. For the latter, for example, consider a binary prospect x_E0, with $u(0)$ assumed to be 0. Eliciting the Certainty-Equivalent (CE) of this prospect and using equation (2.4), we have $v(CE) = w_Q\big(p(E)\big)v(x)$. The decision weight of event E can then be obtained by $w_Q\big(p(E)\big) = v(CE)/v(x)$.

2.5 Measuring the Weighting Function for Risk

Although we are interested in the plausibility and feasibility of the Neo-Lite method for both choice under risk and ambiguity, as a start, we examine the case of risk in this study, mainly because risk is the simplest experimental environment we could construct to test our method. This is an efficient way to obtain an impression of whether the non-fitting nature of the Neo-Lite method works. Application of this method to choice under ambiguity may deserve exclusive research beyond this one.

2.5.1 Experimental Design

The aim of this experiment is to generate a dataset which allows for: (i) the application of both the Neo-Lite method and standard parametric fitting to measure subjects' probability weighting function for risk; (ii) the comparison and discrimination between the two classes of probability weighting functions. Our experiment has two parts to serve each of these purposes.

The Measurement Set

The Measurement Set refers to the dataset for the first part of the experiment. In the first part we elicit the Certainty-Equivalents of 16 risky prospects x_p0, with $x = 4$, 8,

12, 16 and p = 0.2, 0.4, 0.6, 0.8. All four x and all four probabilities are used. We selected four probability levels 0.2, 0.4, 0.6, and 0.8 in order not only to apply the Neo-Lite method, but also to do a robustness check. As has been explained, in the case of risk, the minimum number of decision weights (or probabilities) needed for the Neo-Lite method is three. When there are four different probabilities, there is more than one way to specify the four events $A, B, A \cup B$, and $\bar{A}$. It is unclear whether different specifications yield different measurements of the weighting function, and having at least four probability entries enables an investigation of this question. For simplicity, we didn't have more than four probability entries. Having shown the economy of the Neo-Lite method in data requirements, we choose to apply both our method and parametric fitting to the same data (Measurement Set) in order to make a clean comparison between the two methods.

We introduced some variations of the outcome x because we intend to use the semi-parametric method (Abdellaoui et al. 2008, 2011) to estimate the utility function of money. This method allows for an independent measurement of the utility function, i.e. measuring the utility function without imposing any restriction on the probability weighting function. With this benefit, we can hold the utility function constant for each subject when applying different methods to measure the weighting function. This is important for a clean comparison of the weighting function measurements. More details on the semi-parametric method are provided in the data analysis section.

The 16 CEs are elicited in a randomized order for each subject using choice-list tasks. In our experiment, a choice-list task consists of two choice lists: a basic list and a zoom-in list. Figure 2.2 shows the experimental layout of an example basic list. The upper table shows the gamble to be evaluated. In the experiment the risky prospects were called "gambles". The gambles were presented to subjects in a table, the first row of which contains possible risky events and the second row of which contains the corresponding outcomes of each of the events. Numbers representing the risky event will be explained later in this section. The lower table in this figure is a

The Current Gamble: Gamble 1

From 1 to 12	From 13 to 20
£8.00	£0.00

BASIC LIST 1

	Left Option	Your Choice	Right Option
Choice 1	Gamble 1	Left ○ ○ Right	£0.00
Choice 2	Gamble 1	Left ○ ○ Right	£1.00
Choice 3	Gamble 1	Left ○ ○ Right	£2.00
Choice 4	Gamble 1	Left ○ ○ Right	£3.00
Choice 5	Gamble 1	Left ○ ○ Right	£4.00
Choice 6	Gamble 1	Left ○ ○ Right	£5.00
Choice 7	Gamble 1	Left ○ ○ Right	£6.00
Choice 8	Gamble 1	Left ○ ○ Right	£7.00
Choice 9	Gamble 1	Left ○ ○ Right	£8.00

Figure 2.2: A sample basic choice list

The Current Gamble: Gamble 1

From 1 to 12	From 13 to 20
£8.00	£0.00

ZOOM-IN LIST 1

	Left Option	Your Choice	Right Option
Choice 1	Gamble 1	Left ● ○ Right	£4.00
Choice 2	Gamble 1	Left ○ ○ Right	£4.10
Choice 3	Gamble 1	Left ○ ○ Right	£4.20
Choice 4	Gamble 1	Left ○ ○ Right	£4.30
Choice 5	Gamble 1	Left ○ ○ Right	£4.40
Choice 6	Gamble 1	Left ○ ○ Right	£4.50
Choice 7	Gamble 1	Left ○ ○ Right	£4.60
Choice 8	Gamble 1	Left ○ ○ Right	£4.70
Choice 9	Gamble 1	Left ○ ○ Right	£4.80
Choice 10	Gamble 1	Left ○ ○ Right	£4.90
Choice 11	Gamble 1	Left ○ ● Right	£5.00

Figure 2.3: A sample zoom-in choice list

standard choice list, with the non-risky options (right options) being sure amounts of money from £0 to the gamble prize (£8 in this case) in £1 intervals. As usual, only one switch from choosing left to choosing right is allowed. After the basic list is completed, a zoom-in list will be generated according to the decision maker's choices in the basic list. If the decision maker switches, for instance, between £4 and £5 (i.e. if the gamble is preferred to £4 or less, and £5 or more is preferred to the gamble) in the basic list in Figure 2.2, a zoom-in list as shown in Figure 2.3 will be generated. The non-risky options in the zoom-in list then starts with £4 and end with £5 in steps of £0.10. We introduce zoom-in lists mainly to elicit CEs at greater precision..

The Prediction Set

In the second part of the experiment, subjects complete 96 binary choice questions in randomized orders (see Appendix 2.1 for more details). These choices are referred to as the Prediction Set because we intend to use the measurements of the probability weighting function obtained from the Measurement Set to predict these choices and compare the predictive performance of different functional classes. Figure 2.4 shows the experimental layout of an example binary choice question. For each such question, subjects only have to click either the left or the right button to indicate their choice of the left or the right gamble. Each gamble is either a two-outcome risky prospect or a three-outcome risky prospect, with the outcomes of the gambles being respectively two or three of £6, £10, and £14. Gambles are presented in the same way to subjects as in the first part of the experiment.

In order to distinguish the predictive performance of the Neo-additive weighting function and nonlinear weighting functions, we construct binary choice questions which have two features. First, for each pair of gambles in each choice question, the ratio of the expected payoff of the left gamble to the expected payoff of the right

Figure 2.4: An example binary choice question

gamble is within the range [0.5, 2]. This is to guarantee that no choice is too "easy" (or no gamble is too "obviously" better than its alternative). Second, probabilities involved in these choice questions range from 0.05 to 0.95, in steps of 0.05 for probabilities smaller than 0.2 or larger than 0.8, and in steps of 0.1 for probabilities between 0.2 and 0.8. If the choice questions of the Prediction Set are to be visualized in a Machina (1992)-type unit probability triangle, the segments each of which represents a pair of gambles will symmetrically spread over the whole triangle (see figures in Appendix 2.1).

Due to the second feature, there is higher density for choices involving extreme probabilities than those involving only moderate probabilities. Specifically, 43% of the 96 choice questions involve only probabilities between 0.2 and 0.8 (including 0.2 and 0.8). We call these choices Prediction Subset A for the convenience of later uses (see Figure 2.5 below). For the rest choice questions, at least one of the two gambles involve probabilities smaller than 0.2 or larger than 0.8. We call these choices Prediction Subset B. Since the nonlinear and Neo-additive weighting functions are more likely to differ at two ends of the probability domain than at middle parts, this design may help reveal any potential difference between the two types of weighting functions.
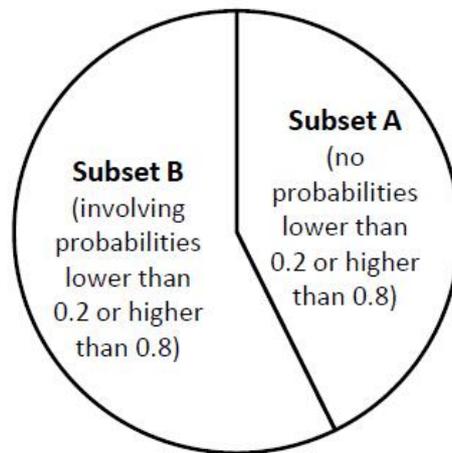
Figure 2.5: Division of the full prediction set

Chance Device and Incentive System

We use a bag filled with 20 numbered balls as the chance device to resolve gambles in this experiment. These numbered balls are from 1 to 20, with each number appearing exactly once. In other words, when one ball is to be drawn from this bag, each number from 1 to 20 is equally likely to be drawn. As shown in Figure 2.2, 2.3 and 2.4, in the experiment, each gamble is presented in the form of a table listing the risky events and the consequences of each event. These risky events relates to which numbered ball is to be drawn from this set of 20 balls. For example, the event "from 1 to 12" refers to "if the number drawn is between 1 and 12", thus denoting a probability of 0.6. We choose exactly twenty numbers because this allows us to have probabilities as small as 0.05 (e.g. "from 1 to 1") and as large as 0.95 (e.g. "from 1 to 19").

We implement a general form of the random incentive system to incentivize subjects. At the end of the experiment, one choice is selected for real for each subject. One ball is then randomly drawn from these 20 balls to determine the results of any gamble chosen by each subject in their selected "real choice". The "real choice" is selected in the following way. For each subject, the computer first randomly selects either the first part or the second part of the experiment with equal probabilities. If

Part 2 is selected, one of the binary choices is randomly selected for real payment. If Part 1 is selected, one of the basic choice lists is randomly selected, and one choice (i.e. one row) in that list is randomly selected. The option chosen for that row determines the real payment, unless the row selected is one of the two 'switch-point' rows in that list. If the latter happens, one of the rows in the corresponding zoom-in choice list will be randomly selected to be the final choice for real payment. This system guarantees that any single choice the subjects make during this experiment could be the one selected for real payment. Since subjects are told that they will receive what they have chosen for the selected 'real choice' as their payoff, this system ensures incentive compatibility.

The experiment was programed and conducted using z-Tree (Fischbacher, 2007) in the CeDEx laboratory at the University of Nottingham. 48 university students were recruited as subjects via the ORSEE system (Greiner, 2015). The experiment lasts approximately 70 minutes. The average payoff for each subject is £12 (including a show-up fee of £3).

2.5.2 Data Analysis

We use the Measurement Set to measure the probability weighting function for risk. Three probability weighting functions are measured, with the second and third belonging to the non-linear class:

$w(p) = \mu + \gamma p$ (Neo-additive, Chateauneuf et al. 2007)

$w(p) = exp(-\gamma(-ln(p))^{\mu})$ (Prelec, 1998)

$w(p) = \gamma p^{\mu} \big/ (\gamma p^{\mu} + (1-p)^{\mu})$ (Goldstein & Einhorn, 1987)

We use semi-parametric estimation to obtain decision weights first, then apply standard parametric fitting for non-linear probability weighting functions and both the Neo-Lite method and the standard parametric fitting for Neo-additive weighting

function. For both methods, we use equation (2.4) and assume the power utility function $v(x) = x^\alpha$.

The reason we also fit the Neo-additive weighting function is to allow for a decomposition of the two-dimensional difference between the Neo-Lite method and the standard parametric fitting (with a nonlinear weighting function). One dimension concerns the measurement techniques (fitting vs. non-fitting) and the other dimension concerns the weighting function forms (nonlinear vs. linear). Prelec's two parameter function and the Goldstein & Einhorn's are chosen as representatives of the non-linear family not only because they are widely used in parametric fitting but also because their parameters have very similar interpretations to those in the Neo-additive function, with μ measuring the degree of elevation (pessimism) and γ measuring the degree of curvature (likelihood-insensitivity). Since it is not clear which of the nonlinear weighting functions fit our data better, we choose to use both and pick up the better one as the benchmark for comparison. We did the following analysis for each subject.

Step 1: *Utility Function* We fitted utility function using the semi-parametric method of Abdellaoui et al. (2008). The power utility function $v(x) = x^\alpha$ is assumed because it is simple, widely used, and often gives a better fit than alternative families (Wakker, 2008). Specifically, we fitted equation (2.4), taking the decision weight $w(p)$ of the outcome-relevant event as an extra parameter. Fitting was done using nonlinear least-square estimation with CE as the dependent variable. The utility was estimated four times at different probability levels, 0.2, 0.4, 0.6, and 0.8. For example, we first estimated the utility parameter, taking $w(0.2)$ as an extra parameter. Then we re-estimate the utility parameter, taking $w(0.4)$ as the extra parameter, etc. In this way we obtained four utility parameter estimates and a full set of estimated decision weights. We tested the equality of the utility parameters estimated with different decision weights and found no significant difference among these utility parameters (details provided in Appendix 2.2). Therefore we take the average of these values as the final utility parameter measurement which we report in the next section.

Table 2.1: Possible specifications of the four events for the Neo-Lite method

Specification	$p(A)$	$p(B)$	$p(A$	$p(\bar{A})$
1	0.2	0.2	0.4	0.8
2	0.2	0.4	0.6	0.8
3	0.2	0.6	0.8	0.8
4	0.4	0.4	0.8	0.6

Note: $p(.)$ denotes probability.

Step 2: *Weighting Function* With the estimated decision weights from the first step, we fitted the probability weighting function parameters using nonlinear least-square estimation. With the same set of decision weights, we also applied the Neo-Lite method. We also did robustness check with respect to the Neo-Lite method by testing the equality of the parameter sets (μ, γ) measured under all the possible different specifications. For risk, each specification assigns a probability to each of the four events $A, B, A \cup B,$ and $\bar{A}$ (see Table 2.1). Under each specification, one set of weighting function parameters can be measured. Again, no significant difference was found between these measurements (more details provided in Appendix 2.3). We therefore take the average of these measurements as the final weighting function parameter measurement (for the Neo-Lite method).

Step 3: Predictive Test Utility functions and probability weighting functions obtained in the first two steps are used to predict the 96 binary choices in the Prediction Set. The predictive success rates, the proportion of correctly predicted choices, are calculated. This indicator of predictive performance does not take into account any stochastic component of decision making, in other words, the measured model can either predict correctly or incorrectly, and the more a model correctly predicts the better the model is.

Due to the non-fitting nature of our method, incorporating decision error analysis in the measurement (and thus in the prediction test) is not possible.

Therefore, all models we fitted are deterministic, with no extra decision error parameter estimated, and the measured models are used deterministically in prediction tests. For the same reason, we did not choose any probabilistic indicators of predictive performance (such as the predicted log-likelihood used by Kothiyal et al. 2014) that are not comparable with the predictive success rate. Another reason for using a "deterministic" indicator is that it provides information about how useful these models are in predicting actual choices.

2.5.3 Experimental Results

We present results for 47 of the 48 subjects, with Subject 7 excluded from the analysis for that he exhibited the same CE for all 16 prospects in the first part of the experiment and always chose left in second part[8]. We think this subject was clearly using extremely simple decision rules, which is not helpful for our intended analysis, though including this subject would not have changed the qualitative results presented in this section.

Measurements

For the power utility function $v(x) = x^\alpha$, the median utility parameter estimated using the semi-parametric method is $\alpha = 0.870$, significantly smaller than 1 (Wilcoxon sign-rank test, p-value 0.000). Median parameter values are considered instead of mean values because there is some skewness in the distribution of the utility estimates (see Figure 2.6).

---

[8] Actually we looked at each subject's choices and excluded subjects whose choices in the experiment are unlikely a reflection of their attitudes towards risk and incentives in real life. For this reason, we excluded subjects with either or both of the two features: (1) a zero or negative utility parameter α; (2) a probability weighting function part or all of which is downward-sloping. Luckily, no other subject other than Subject 7 needs to be excluded.
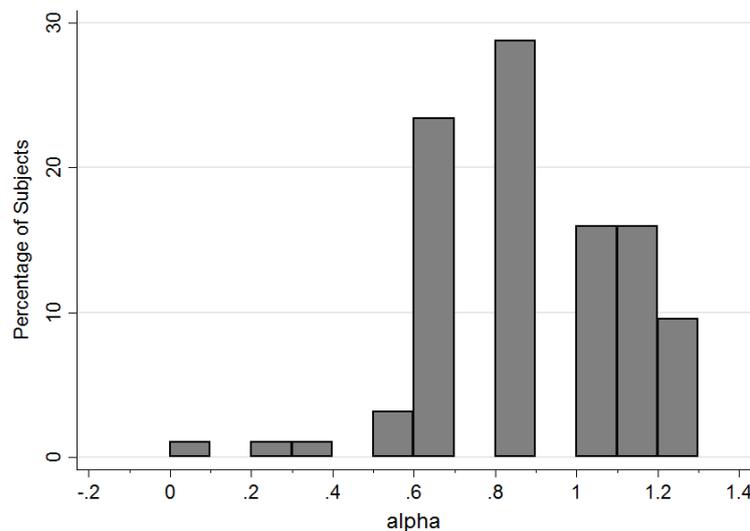
Figure 2.6: Distributions of the measured utility parameter

Figure 2.6 shows the histogram distribution of the measured utility parameter. In general, there is slightly more subjects with concave utility functions than with convex ones. The excluded subject corresponds to the short bar to the most left in this figure, having an almost flat utility function with $\alpha = 0.017$, which is not significantly different from 0 according to the nonlinear Wald-test (p-value 0.324).

The four decision weights estimated using the semi-parametric method are summarized in Table 2.2. Significant deviation from the Expected-Utility theory is found, as shown in the fifth column of Table 2.2.

Table 2.2: Decision weights estimated with semi-parametric fitting

Decision Weight	Mean	Median	Interquartile	t-test p-value $(w(p) = p)$
$w(0.2)$	0.379	0.355	[0.250, 0.481]	0.000
$w(0.4)$	0.508	0.495	[0.400, 0.630]	0.000
$w(0.6)$	0.622	0.624	[0.524, 0.727]	0.048
$w(0.8)$	0.748	0.761	[0.656, 0.856]	0.000

For the Neo-Lite methods, we first measured μ according to equation (2.8),

and measured the pessimism/optimism index according to equation (2.10). The median intercept $\mu = 0.26$ and the median pessimism/optimism index $(1 - \gamma - 2\mu) = -0.11$. Then the slope γ is solved from these two results. The median measurement is $\gamma = 0.59$. The results of all the parameter measurements are listed in Table 2.3.

Table 2.3: Median weighting function parameter measurements

	Prelec $w(p) = exp(-\gamma(-\ln(p))^{\mu})$	**G&E** $w(p) = \gamma p^{\mu} \big/ (\gamma p^{\mu} + (1-p)^{\mu})$	**Neo** $w(p) = \mu + \gamma p$	**Neo-Lite** $w(p) = \mu + \gamma p$
μ	0.67	0.61	0.26	0.26
γ	0.75	1.30	0.58	0.59

In Table 2.3, "Neo-Lite" denotes the Neo-additive weighting function measured using Neo-Lite method; "Neo", "Prelec" and "G&E" denote respectively the Neo-additive weighting function, Prelec (1998)'s weighting function and Goldstein & Einhorn (1987)'s weighting function, all measured using parametric fitting. We will continue to use these meanings and denotations for the rest of the chapter. Again, due to skewness in the distribution of the parameters, we report the median weighting functions. It is worth noticing that although the parameters of the Prelec function and the G&E function have similar interpretations, their parameter values (especially of γ) are not comparable. The difference between the γ of "Prelec" and of "G&E" actually entails a similarity in their functional shapes, as is shown in the figure below.
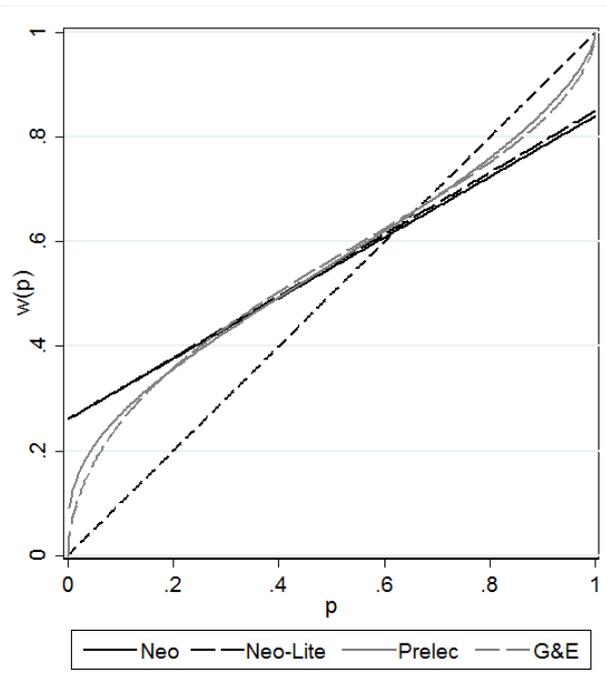
Figure 2.7: Median Probability Weighting Functions

In Figure 2.7 we plot the four weighting functions listed in Table 2.3. It is not surprising that all non-linear weighting functions exhibit the common inverse-S pattern. With hindsight it is also not so surprising that for probabilities between 0.2 and 0.8, the fitted weighting functions almost overlap, since we have only four observations of decision weights that are between 0.2 and 0.8. However, it is surprising that the median "Neo-Lite" weighting function is almost exactly the same as the "Neo" weighting function, given that completed different measurement techniques (non-fitting and fitting procedures) are used. In order to further compare the two techniques, we looked at the parameter distributions of the "Neo-Lite" weighting function and the "Neo" weighting function. The left histogram of Figure 2.8 shows the distributions of measured μ (the intercept) and the right histogram shows the distributions of measured γ (the slope).
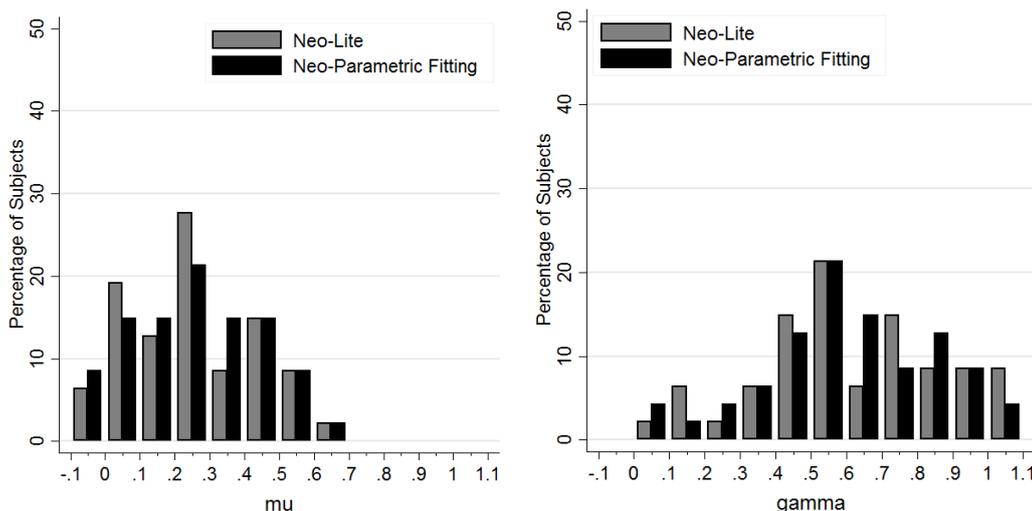
Figure 2.8: Distributions of parameters of the Neo-additive weighting function

For μ, although negative values are not allowed in theory, we found a few subjects having a negative μ. This is not strange if one has an S-shaped weighting function but only the middle part of that function is measured with a linear functional form. Therefore we didn't exclude subjects with negative μ because they may represent a certain, though not common, type of subjects (e.g. having an S-shaped weighting function).

We used three tests to inspect whether the distributions of μ and γ of the Neo-additive function measured using Neo-Lite method and using parametric fitting are the same: the t-test (for equality of means), the Wilcoxon sign-rank test (for the equality of medians), and the Mann-Whitney two-sample test (for the equality of the entire distributions). The result is that, for both μ and γ, there is no significant difference (at 5% level) of the means, medians, or the whole distributions between the two methods. The p-values of the three tests are respectively 0.541, 0.612, 0.979 for μ, and 0.419, 0.216, 0.889 for γ.

We also compared the resulting distributions of different types of subjects from different measurement methods. Subjects are classified according to their attitudes towards probabilities (i.e. how they weight probabilities)[9]. Each subject is classified

---

[9] Since the measured median utility function is significantly different from linearity, the probability weighting function only reflects the decision maker's attitude towards probabilities rather than attitude towards risk.

as one of the four types: inverse-S (for subjects overweighting small probabilities and underweighting large probabilities), S-shape (for subjects underweighting small probabilities and overweighting large probabilities), overweighting (for subjects overweighting all probabilities), and underweighting (for subjects underweighting all probabilities). Details of how subjects are classified are provided in Appendix 2.4. Figure 2.9 reports the (stack) distributions of the numbers of the four types of subjects for each of the four weighting functions "Neo", "Neo-Lite", "Prelec", and "G&E".
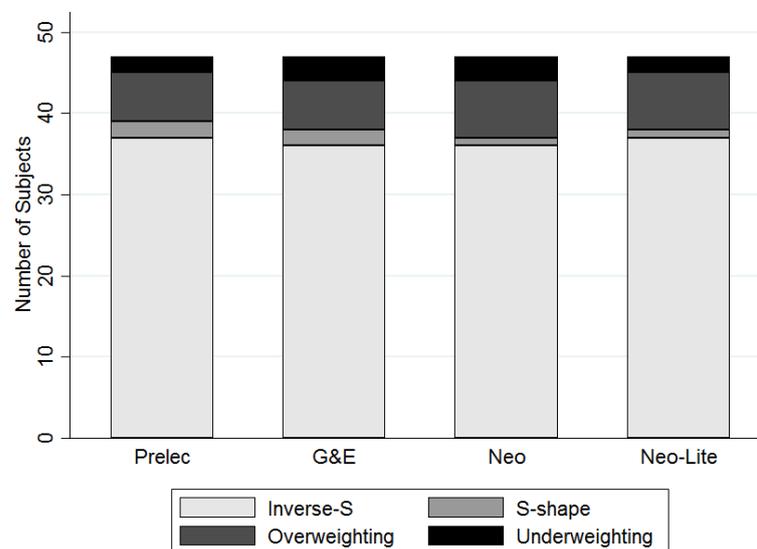
Figure 2.9: Distributions of different types of subjects according to their attitude towards probability

Each bar in Figure 2.9 is a stack of (from bottom to top) the number of the inverse-S type, the S-shape type, the overweighting type and the underweighting type, with numbers of all four types summing up to 47. Not surprisingly, regardless of the measurement methods or weighting function models, most of the subjects were classified as having an inverse-S shaped weighting function, as denoted by the lightest color in the figure. More importantly, according to the Mann-Whitney test of equality of distributions, all four bars have the same distribution of subjects with

respect to the four types of probability weighting[10].

These results suggest that the non-fitting technique of the new method seems to be a good substitute for parametric fitting. In other words, when the weighting function is restricted to be Neo-additive, the Neo-Lite method and the parametric fitting method generate essentially the same measurements. Indeed, we couldn't find any statistically significant difference in all these three aspects of the two measurements. However, the Neo-Lite method seems to be more efficient in this case than parametric fitting of the decision weighting function, because with the former, parameters could simply be calculated from decision weights without extra work.

Out-of-Sample Predictive Performance

We expected the predictive performance test to help us distinguish and discriminate between the four probability weighting functions, especially between "Neo-Lite" and "Neo", and between "Neo" and the two nonlinear weighting functions. The former comparison relates to the comparison between data fitting technique and the non-fitting technique of our method, while the latter relates to the comparison between the descriptive and predictive power of the Neo-additive (linear) weighting function and the nonlinear weighting functions. Since we have shown that the "Neo-Lite" functions are not significantly different from the "Neo" functions, we wouldn't expect their average predictive performances differ significantly (though the performance distribution across subjects may differ). However, by the nature of our design of the prediction set, we do expect that either the nonlinear weighting functions performs significantly better than the Neo-additive functions, or the opposite.

We calculate the predictive success rate for each subject with each type of probability weighting function, and report the mean success rate over all subjects in Table 2.4. The predictive success rate is calculated as the proportion of correctly

---

[10] Since the Mann-Whitney test can only test the equality of two samples, we did this test for each two of the four weighting functions, and therefore did six tests. The minimum p-value obtained among the six tests is 0.743.

predicted choices in all choices considered. We use Predictive Test 1 to refer to the test in which we use the (probability weighting functions obtained from the) Measurement Set to predict (choices in) the Prediction Set. In general, predictive success rates are low when the Prediction Set is considered as the predictive target.

Table 2.4: Mean Predictive Success Rate (for the whole Prediction Set)

Prelec	G&E	Neo	Neo-Lite
52.3%	51.1%	52.0%	51.7%

Although, not surprisingly, there is no significant difference between the performance of "Neo" and "Neo-Lite", it is unexpected somehow that there is also no significant difference between the nonlinear weighting function class and the Neo-additive class. The only significant difference lies between "Prelec" and "G&E" (t-test, p-value 0.005). It seems that our prediction set does not achieve the purpose of distinguishing the nonlinear weighting functions from the Neo-additive ones. We speculate that this is because the four measured functions perform equally bad given that this predictive test is a very strong out-of-sample test, with measurements of the probability weighting function from four probabilities (0.2, 0.4, 0.6, 0.8) to predict binary choices involving probabilities from 0.05 to 0.95. In other words, there might be a common failure of the four functions to predict choices involving extreme probabilities, and this common failure may have largely reduced the performance gap between the two classes which would otherwise be salient. Therefore, we changed the predictive target from the full prediction set to its subsets (Subset A and B as mentioned earlier) and examine the predictive performances.

Table 2.5 shows the mean predictive success rate for the four weighting functions when predicting Prediction Subset A (where only probabilities between 0.2 and 0.8 are involved) and Prediction Subset B (where probabilities lower than 0.2 or higher than 0.8 are involved). There are three key messages from Table 2.5. First, for Prediction Subset A, the success rates are significantly improved whereas for

Prediction Subset B, the opposite happens. This is again expected because in the Measurement Set, only probabilities between 0.2 and 0.8 are involved. In other words, the two ends of the weighting functions are not measured. Secondly, for Prediction Subset B, the predictive success rates even dip below 50%, indicating a systematic deviation of subjects' attitudes towards probability from what has been elicited from the Measurement Set. Thirdly, whereas for Prediction Subset B there is no significant difference between the Neo-additive weighting functions and the nonlinear weighting functions, for Subset A, the nonlinear class performs significantly better than the Neo-additive class[11]. Subset B seems to be where the common failure lies.

Table 2.5: Mean Predictive Success Rate (for subsets of the Prediction Set)

Predictive Target	**Prelec**	**G&E**	**Neo**	**Neo-Lite**
Prediction Subset A (non-extreme probabilities)	58.0%	57.9%	56.9%	56.5%
Prediction Subset B (extreme probabilities)	47.5%	48.0%	48.4%	48.2%

In-Sample Predictive Exercise

We intended to discriminate between the Neo-additive weighting functions and the nonlinear weighting functions, but obviously the result of the out-of-sample test is enough for us to claim support for either of the two classes. Although for moderate probabilities, the nonlinear class performs better than the Neo-additive class, the absolute difference in the mean predictive success rates is less than 2%. Moreover, we failed to discriminate between the nonlinear weighting functions from the

---

[11] We did t-tests for the equality of the mean predictive success rates between each pair of the linear-nonlinear weighting functions. The p-values for the null hypotheses "Prelec"= "Neo", "Prelec"= "Neo-Lite", "G&E"= "Neo", and "G&E"= "Neo-Lite" are respectively 0.015, 0.031, 0.019, and 0.037.

Neo-additive ones at the probability domain (i.e. Subset B) where both classes of weighting functions fail to predict.

It may be argued that our out-of-sample predictive test is perhaps too hard for the CPT model we work with, in light of the well-known preference reversal phenomenon (e.g. Lichtenstein & Slovic, 1971; Grether & Plott, 1979). For this reason, we speculate that performing in-sample tests, or using the binary choices as both a measurement set and a prediction set, would generally improve the predictive success rates of both classes of weighting functions, and this may also make the differences of performances distinguishable between the two classes. So we fitted the probability weighting functions "Prelec", "G&E" and "Neo" with binary choices from Prediction Subset A, using the Maximum Likelihood Estimation introduced by Harrison (2008)[12]. We chose Prediction Subset A (instead of the full Prediction Set) as the new measurement set in order to make it more comparable to the original one because now both involve only probabilities between 0.2 and 0.8. We didn't apply the Neo-Lite method in this case because it is much more complicated to obtain the decision weights first (than to fit the utility and weighting function parameters directly). However, this does not matter because the purpose of this additional analysis is to discriminate between the Neo-additive class and the nonlinear class of weighting functions. In other words, we need only to compare "Neo" with "Prelec" or "G&E".

We simultaneously fitted a power utility function with each of the three weighting functions (estimation results are shown in Appendix 2.5). The measured utility and weighting functions were used to predict the full Prediction Set, within which about 60% of choices involved probabilities lower than 0.2 or higher than 0.8. Table 2.6 reports the results.

In this case we used a subset of the Prediction Set to predict the full set and the two complementary subsets. For all three predictive targets, the nonlinear weighting functions now performs significantly better than the fitted "Neo" function, with an

---

[12] Details of these estimations can be found in Appendix 2.5.

absolute gap from approximately 3% to 9%[13]. This suggests that the nonlinear probability weighting functions seem to have more predictive (and hence descriptive) power than the Neo-additive weighting functions. In addition, Table 2.6 also shows that, again, choices involving extreme probabilities (Prediction Subset B) are harder to predict than choices involving no extreme probabilities.

Table 2.6: Mean Benchmark Predictive Success Rate

Predictive Target	**Prelec**	**G&E**	**Neo**
Full Prediction Set	66.7%	69.2%	62.6%
Prediction Subset A (non-extreme probabilities)	65.7%	71.2%	64.3%
Prediction Subset B (extreme probabilities)	67.0%	68.3%	62.3%

Note: Prediction subset A is also the measurement set in this case.

It now seems that although the Neo-Lite method works almost as well as parametric fitting when the Neo-additive weighting function is used, as shown by our evidence that that the Neo-Lite method generate almost exactly the same measurements of the Neo-additive weighting function as those generated by parametric fitting (with the robustness check of our method indicating its capability to exploiting more data which potentially favors parametric fitting), our method may not be a good substitute of parametric fitting because, in general, the Neo-additive weighting function seems to have less predictive power than the nonlinear class of weighting functions, especially when extreme probabilities are involved.

So far we have done six predictive exercises. For convenience of analysis, the six exercises are named from Predictive Test 1 to 6 as shown in the table below. Strictly speaking, only Prediction Test 5 is an in-sample test, and all others are out-of-sample tests. The results for Test 1 is shown in Table 2.4, results for Test 2 and 3 are shown in Table 2.5, and results for Test 4, 5, and 6 are shown in Table 2.6.

---

[13] For each predictive target, we did t-tests for the equality of the mean predictive success rates between each pair of the three functions. All p-values are 0.000. This time, the "G&E" function performs significantly better than "Prelec" and "Neo".

CHAPTER 2

In general, we can see that the performance of Test 4, 5 and 6 is much better than Test 1, 2, and 3. In other words, we used two different measurement sets (one based on CEs and the other based on binary choices) to measure subjects' utility and probability weighting functions, and the two measurements have substantially different predictive performances in predicting the same three sets of choices. This is somewhat unexpected because the two measurement sets involve the same range of probabilities (from 0.2 to 0.8).

Table 2.7: The six predictive tests

	Probability weighting functions measured from	Predictive target
Predictive Test 1	the Measurement Set	Prediction Set
Predictive Test 2	the Measurement Set	Prediction Subset A
Predictive Test 3	the Measurement Set	Prediction Subset B
Predictive Test 4	Prediction Subset A	Prediction Set
Predictive Test 5	Prediction Subset A	Prediction Subset A
Predictive Test 6	Prediction Subset A	Prediction Subset B

Not surprisingly, the in-sample predictive success rates (Test 5) are significantly better than the corresponding out-of-sample success rates (Test 2). However, the facts that success rates in Test 4 are much higher than in Test 1, and that success rates in Test 6 are much higher than in Test 3 suggest that there might be some crucial difference between probability weighting functions measured using CE data and those measured using choice data. This is confirmed by Figure 2.10, which shows the probability weighting function we estimated from choices of Prediction Subset A. These weighting functions are quite different from those we obtained using the Measurement Set (i.e. CE data). The estimated weighting functions exhibit an uncommon shape: almost S-shaped for the nonlinear weighting functions and universally underweighting for the Neo-additive function.
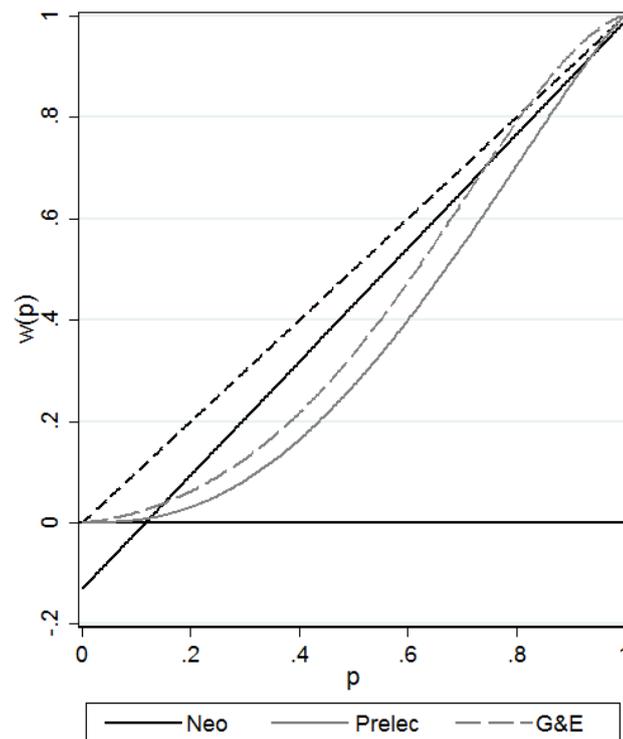
Figure 2.10: Median probability weighting functions estimated from Prediction Subset A

This finding has reminded us of the famous preference reversal phenomenon (see e.g. Seidl, 2002 for a survey), that is inconsistency between preference revealed from binary choices tasks and from evaluation tasks. Various explanations for this phenomenon have been studied (e.g. Loomes & Sugden, 1983; Loomes et al., 1989; Cubitt et al., 2004). However, our finding that different types of tasks of decision-making under risk reveal completely different probability weighting functions (inverse-S vs. S-shaped) brings up new interesting questions. For example, is preference reversal (and factors contributing to preference reversal) a cause for or a result of individuals having different weighting functions for different types of tasks? If the latter is true, what could then explain why people have different weighting functions for choice tasks and for evaluation tasks? These questions may deserve future research.

2.6 Conclusions

We have introduced a non-fitting method, based on the Neo-additive parametric form and the subadditivity properties of the probability weighting function, to measure the probability weighting function for choice under uncertainty. For risk, the advantage of this method is mainly efficiency of the data analysis process. For ambiguity, the main advantage is that no elicitation of the subjective probabilities is needed. In this chapter, we did an experiment of choice under risk to test our method against standard parametric fitting method, and to test the non-linear class of probability weighting function against the Neo-additive weighting function.

We have shown with our experiment that, in combination with a method that elicits decision weights, the Neo-Lite method provides a most simple and quick way to obtain measurements of the Neo-additive weighting function. These measurements do not differ from the Neo-additive weighting function measured using standard parametric fitting. However, as our predictive tests show, the non-linear class of probability weighting functions generally performs better than the Neo-additive function, particularly when we use choice data to predict choices. The usefulness of our method seems to be bounded by the natural limitations of the linear weighting functions to describe choices involving extreme probabilities. In other words, the Neo-Lite method works well whenever the Neo-additive weighting function is considered appropriate for use. It is not clear how our method works for ambiguity, but for risk, we generally suggest using this method as a quick and convenient way to obtain some pre-analysis results or to get a sense of the degree of pessimism and likelihood insensitivity of the decision maker, than as a rigorous empirical method in formal decision analysis.

In addition, our finding that people exhibit dramatically different shapes of probability weighting raises three important questions. The first is to what extent is the inverse-S shaped probability weighting function representative of individual risk preferences. The second is how we should interpret a probability weighting function

that is sensitive to the type of task used to elicit preferences. The third question is whether there are more fundamental regularities underlying decision-making that can explain this sensitivity. The answers to these questions may provide insights for how CPT can be extended or modified, or how future theories can be developed to accommodate violations of EUT as well as the preference reversal phenomenon.

Appendices for Chapter 2

Appendix 2.1 More details of the Prediction Set

Table 2.8: The list of all the binary choice questions in the Prediction Set of the experiment

Question	Left Option	Right Option
1	(£14, 0.9 ; £10, 0.1)	(£14, 0.95 ; £6, 0.05)
2	(£14, 0.85 ; £10, 0.15)	(£14, 0.95 ; £6, 0.05)
3	(£14, 0.85 ; £10, 0.15)	(£14, 0.9 ; £6, 0.1)
4	(£14, 0.8 ; £10, 0.2)	(£14, 0.9 ; £6, 0.1)
5	(£14, 0.7 ; £10, 0.3)	(£14, 0.9 ; £6, 0.1)
6	(£14, 0.7 ; £10, 0.3)	(£14, 0.85 ; £6, 0.15)
7	(£14, 0.7 ; £10, 0.3)	(£14, 0.8 ; £6, 0.2)
8	(£14, 0.6 ; £10, 0.4)	(£14, 0.85 ; £6, 0.15)
9	(£14, 0.6 ; £10, 0.4)	(£14, 0.8 ; £6, 0.2)
10	(£14, 0.5 ; £10, 0.5)	(£14, 0.8 ; £6, 0.2)
11	(£14, 0.5 ; £10, 0.5)	(£14, 0.7 ; £6, 0.3)
12	(£14, 0.4 ; £10, 0.6)	(£14, 0.8 ; £6, 0.2)
13	(£14, 0.4 ; £10, 0.6)	(£14, 0.7 ; £6, 0.3)
14	(£14, 0.4 ; £10, 0.6)	(£14, 0.6 ; £6, 0.4)
15	(£14, 0.3 ; £10, 0.7)	(£14, 0.7 ; £6, 0.3)
16	(£14, 0.3 ; £10, 0.7)	(£14, 0.6 ; £6, 0.4)
17	(£14, 0.2 ; £10, 0.8)	(£14, 0.7 ; £6, 0.3)
18	(£14, 0.2 ; £10, 0.8)	(£14, 0.6 ; £6, 0.4)
19	(£14, 0.2 ; £10, 0.8)	(£14, 0.5 ; £6, 0.5)
20	(£14, 0.15 ; £10, 0.85)	(£14, 0.7 ; £6, 0.3)
21	(£14, 0.15 ; £10, 0.85)	(£14, 0.6 ; £6, 0.4)
22	(£14, 0.15 ; £10, 0.85)	(£14, 0.5 ; £6, 0.5)
23	(£14, 0.6 ; £10, 0.4)	(£14, 0.6 ; £6, 0.4)
24	(£14, 0.1 ; £10, 0.9)	(£14, 0.6 ; £6, 0.4)
25	(£14, 0.1 ; £10, 0.9)	(£14, 0.5 ; £6, 0.5)
26	(£14, 0.1 ; £10, 0.9)	(£14, 0.4 ; £6, 0.6)
27	(£14, 0.05 ; £10, 0.95)	(£14, 0.6 ; £6, 0.4)
28	(£14, 0.05 ; £10, 0.95)	(£14, 0.5 ; £6, 0.5)
29	(£14, 0.05 ; £10, 0.95)	(£14, 0.4 ; £6, 0.6)
30	(£10, 0.95 ; £6, 0.05)	(£14, 0.4 ; £6, 0.6)
31	(£10, 0.95 ; £6, 0.05)	(£14, 0.5 ; £6, 0.5)
32	(£10, 0.95 ; £6, 0.05)	(£14, 0.6 ; £6, 0.4)

33	(£10, 0.9 ; £6, 0.1)	(£14, 0.6 ; £6, 0.4)
34	(£10, 0.9 ; £6, 0.1)	(£14, 0.5 ; £6, 0.5)
35	(£10, 0.9 ; £6, 0.1)	(£14, 0.4 ; £6, 0.6)
36	(£10, 0.2 ; £6, 0.8)	(£14, 0.25 ; £6, 0.75)
37	(£10, 0.85 ; £6, 0.15)	(£14, 0.5 ; £6, 0.5)
38	(£10, 0.85 ; £6, 0.15)	(£14, 0.4 ; £6, 0.6)
39	(£10, 0.85 ; £6, 0.15)	(£14, 0.3 ; £6, 0.7)
40	(£10, 0.8 ; £6, 0.2)	(£14, 0.5 ; £6, 0.5)
41	(£10, 0.8 ; £6, 0.2)	(£14, 0.4 ; £6, 0.6)
42	(£10, 0.8 ; £6, 0.2)	(£14, 0.3 ; £6, 0.7)
43	(£10, 0.7 ; £6, 0.3)	(£14, 0.4 ; £6, 0.6)
44	(£10, 0.7 ; £6, 0.3)	(£14, 0.3 ; £6, 0.7)
45	(£10, 0.6 ; £6, 0.4)	(£14, 0.4 ; £6, 0.6)
46	(£10, 0.6 ; £6, 0.4)	(£14, 0.3 ; £6, 0.7)
47	(£10, 0.6 ; £6, 0.4)	(£14, 0.2 ; £6, 0.8)
48	(£10, 0.5 ; £6, 0.5)	(£14, 0.3 ; £6, 0.7)
49	(£10, 0.5 ; £6, 0.5)	(£14, 0.2 ; £6, 0.8)
50	(£10, 0.4 ; £6, 0.6)	(£14, 0.2 ; £6, 0.8)
51	(£10, 0.4 ; £6, 0.6)	(£14, 0.15 ; £6, 0.85)
52	(£10, 0.3 ; £6, 0.7)	(£14, 0.2 ; £6, 0.8)
53	(£10, 0.3 ; £6, 0.7)	(£14, 0.15 ; £6, 0.85)
54	(£10, 0.3 ; £6, 0.7)	(£14, 0.1 ; £6, 0.9)
55	(£10, 0.2 ; £6, 0.8)	(£14, 0.1 ; £6, 0.9)
56	(£10, 0.15 ; £6, 0.85)	(£14, 0.1 ; £6, 0.9)
57	(£10, 0.15 ; £6, 0.85)	(£14, 0.05 ; £6, 0.95)
58	(£10, 0.1 ; £6, 0.9)	(£14, 0.05 ; £6, 0.95)
59	(£10, 0.75 ; £6, 0.25)	(£14, 0.5 £10 0.5)
60	(£14, 0.85 ; £10, 0.1 ; £6, 0.05)	(£14, 0.9 ; £6, 0.1)
61	(£14, 0.8 ; £10, 0.15 ; £6, 0.05)	(£14, 0.9 ; £6, 0.1)
62	(£14, 0.8 ; £10, 0.15 ; £6, 0.05)	(£14, 0.85 ; £6, 0.15)
63	(£14, 0.7 ; £10, 0.25 ; £6, 0.05)	(£14, 0.85 ; £6, 0.15)
64	(£14, 0.7 ; £10, 0.25 ; £6, 0.05)	(£14, 0.85 ; £6, 0.15)
65	(£14, 0.8 ; £10, 0.1 ; £6, 0.1)	(£14, 0.8 ; £6, 0.2)
66	(£14, 0.7 ; £10, 0.2 ; £6, 0.1)	(£14, 0.8 ; £6, 0.2)
67	(£14, 0.7 ; £10, 0.15 ; £6, 0.15)	(£14, 0.8 ; £6, 0.2)
68	(£14, 0.6 ; £10, 0.3 ; £6, 0.1)	(£14, 0.7 ; £6, 0.3)
69	(£14, 0.6 ; £10, 0.25 ; £6, 0.15)	(£14, 0.7 ; £6, 0.3)
70	(£14, 0.5 ; £10, 0.4 ; £6, 0.1)	(£14, 0.7 ; £6, 0.3)
71	(£14, 0.5 ; £10, 0.35 ; £6, 0.15)	(£14, 0.7 ; £6, 0.3)
72	(£14, 0.4 ; £10, 0.5 ; £6, 0.1)	(£14, 0.6 ; £6, 0.4)
73	(£14, 0.3 ; £10, 0.65 ; £6, 0.05)	(£14, 0.6 ; £6, 0.4)
74	(£14, 0.25 ; £10, 0.6 ; £6, 0.15)	(£14, 0.6 ; £6, 0.4)
75	(£14, 0.25 ; £10, 0.6 ; £6, 0.15)	(£14, 0.5 ; £6, 0.5)
76	(£14, 0.15 ; £10, 0.6 ; £6, 0.25)	(£14, 0.5 ; £6, 0.5)

77	(£14, 0.15 ; £10, 0.6 ; £6, 0.25)	(£14, 0.4 ; £6, 0.6)
78	(£14, 0.05 ; £10, 0.65 ; £6, 0.3)	(£14, 0.4 ; £6, 0.6)
79	(£14, 0.1 ; £10, 0.5 ; £6, 0.4)	(£14, 0.4 ; £6, 0.6)
80	(£14, 0.15 ; £10, 0.35 ; £6, 0.5)	(£14, 0.3 ; £6, 0.7)
81	(£14, 0.1 ; £10, 0.4 ; £6, 0.5)	(£14, 0.3 ; £6, 0.7)
82	(£14, 0.15 ; £10, 0.25 ; £6, 0.6)	(£14, 0.3 ; £6, 0.7)
83	(£14, 0.1 ; £10, 0.3 ; £6, 0.6)	(£14, 0.3 ; £6, 0.7)
84	(£14, 0.15 ; £10, 0.15 ; £6, 0.7)	(£14, 0.2 ; £6, 0.8)
85	(£14, 0.1 ; £10, 0.2 ; £6, 0.7)	(£14, 0.2 ; £6, 0.8)
86	(£14, 0.05 ; £10, 0.25 ; £6, 0.7)	(£14, 0.2 ; £6, 0.8)
87	(£14, 0.05 ; £10, 0.25 ; £6, 0.7)	(£14, 0.15 ; £6, 0.85)
88	(£14, 0.1 ; £10, 0.1 ; £6, 0.8)	(£14, 0.15 ; £6, 0.85)
89	(£14, 0.05 ; £10, 0.15 ; £6, 0.8)	(£14, 0.15 ; £6, 0.85)
90	(£14, 0.05 ; £10, 0.15 ; £6, 0.8)	(£14, 0.1 ; £6, 0.9)
91	(£14, 0.05 ; £10, 0.1 ; £6, 0.85)	(£14, 0.1 ; £6, 0.9)
92	£10	(£14, 0.6 ; £6, 0.4)
93	£10	(£14, 0.5 ; £6, 0.5)
94	£10	(£14, 0.4 ; £6, 0.6)
95	£10	(£14, 0.2 ; £10, 0.6 ; £6, 0.2)
96	£10	(£14, 0.1 ; £10, 0.8 ; £6, 0.1)

In Figure 2.11, we sketch all the choices in the table above on a unit probability triangle. Each choice question is represented by a segment in the triangle, with each of the two end nodes of a segment representing each of the two gambles in a choice question. The triangle is positioned in a square divided into 20 rows and 20 columns. Each row (column) has a height (width) of 1/20 of the side length of the large square. In other words all three sides of the probability triangle are divided into probability intervals of 0.05. At the top corner is the best outcome £14, at the left bottom corner is the middle outcome £10, and at the right bottom corner is the worst outcome £6. Figure 2.11a sketches the choices with two-outcome gambles and Figure 2.11b shows choices involving three-outcome gambles.
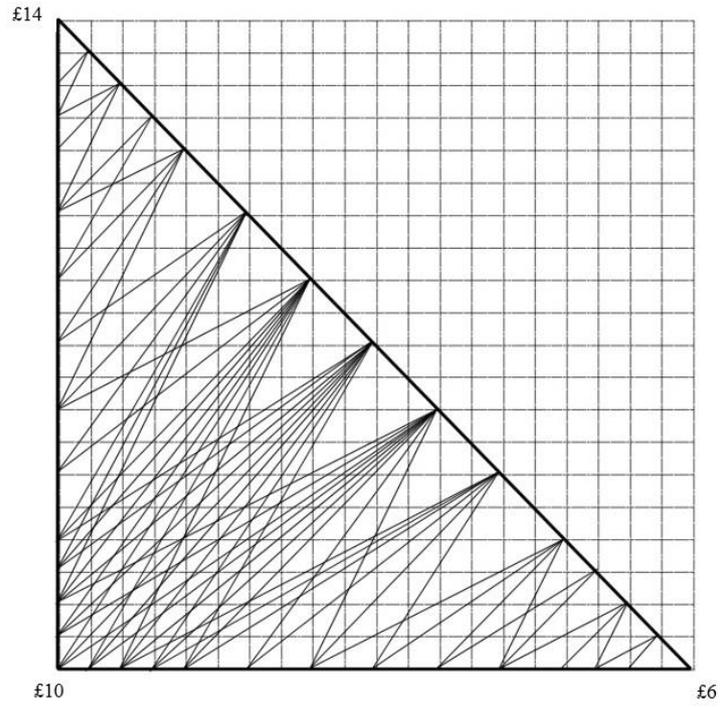
Figure 2.11a: Choice questions of the Prediction Set involving only two-outcome gambles
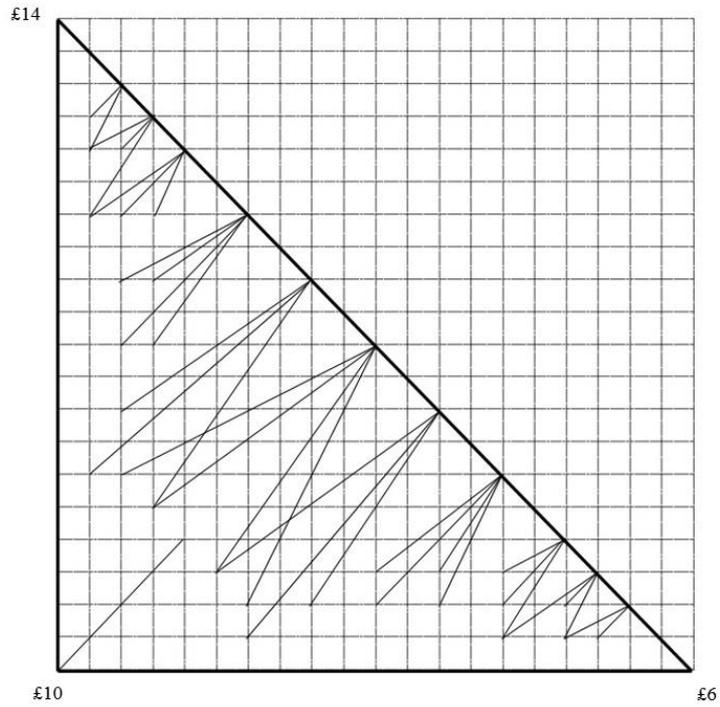
Figure 2.11b: Choice questions of the Prediction Set involving three-outcome gambles

Appendix 2.2 Measurements of utility

To use the semi-parametric method to estimate the utility function and the decision weights, we first divided the Measurement Set into four subsets. Each of the subsets contains four Certainty-Equivalents of gambles having the same probability but different stakes. For each of these subsets of CEs, we estimated the power utility parameter α, taking the corresponding decision weight as an extra parameter (see Table 8). Since in the Measurement Set we have four probabilities 0.2, 0.4, 0.6, and 0.8, we ended up with four utility measurements for each of the probabilities (or decision weights) for each subject.

Table 2.9 Semi-parametric fitting of the utility parameter and decision weights

Data for Estimation	**Parameters Estimated**	**Estimated Utility**	
		Median	**Mean**
CE($4_{0.2}0$), CE($8_{0.2}0$), CE($12_{0.2}0$), CE($16_{0.2}0$)	α, $w(0.2)$	$\alpha = 0.86$	$\alpha = 0.88$
CE($4_{0.4}0$), CE($8_{0.4}0$), CE($12_{0.4}0$), CE($16_{0.4}0$)	α, $w(0.4)$	$\alpha = 0.84$	$\alpha = 0.83$
CE($4_{0.6}0$), CE($8_{0.6}0$), CE($12_{0.6}0$), CE($16_{0.6}0$)	α, $w(0.6)$	$\alpha = 0.86$	$\alpha = 0.85$
CE($4_{0.8}0$), CE($8_{0.8}0$), CE($12_{0.8}0$), CE($16_{0.8}0$)	α, $w(0.8)$	$\alpha = 0.90$	$\alpha = 0.88$

Note: The model estimated is equation (2.4), with power utility $v(x) = x^{\alpha}$. CE($x_p y$) denotes the Certainty-Equivalent of the gamble $x_p y$.

We tested the equality of the four utility measurements at the aggregate level. Specifically, we tested the equality of the median utility parameters (the third column of Table 2.8) using Wilcoxon sign-rank test and found no significant difference any two of the four measurements (the minimum p-value 0.243). We tested the equality of the mean utility parameters (the fourth column of Table 2.8) using t-test and found no significant difference between any two of the four measurements (the minimum p-value 0.435). We also tested the hypothesis that these measurements are drawn from the populations with the same distribution using the Mann-Whitney test, and again no significant difference has been found (the minimum p-value 0.275). Therefore for each subject, we averaged the four utility parameter measurements and reported the averaged median utility parameter (over subjects) and the averaged utility parameter distribution (over subjects) in Section 2.5.3 of the main text.

Appendix 2.3 Robustness checks for Neo-Lite method

After obtaining, for each subject, the four decision weights $w(0.2)$, $w(0.4)$, $w(0.6)$ and $w(0.8)$, we could obtain four potentially different sets of weighting function parameters for each different specifications of the relevant events (see Table 2.9). We therefore did a robustness check for the Neo-Lite method by testing the equality of the four weighting function measurements, as what we have done for the four utility measurements generated using different subsets of the Measurement Set.

At the aggregate level, we again used t-test to test the equality of means, Wilcoxon sign-rank test to test the equality of medians, and Mann-Whitney test to test the equality of the distributions. We found no significant difference between any two of these measurements of μ and γ. Specifically, for μ, the minimum p-values for mean tests, median test, and distribution tests are respectively 0.415, 0.560, and 0.466; for γ, the minimum p-values for mean tests, median test, and distribution tests are respectively 0.412, 0.382, and 0.620. We therefore average the four

measurements of μ and γ for each subject and reported the median (over subjects) of the averaged weighting function parameters in Table 2.3.

Table 2.10: Median weighting function parameters measured using Neo-Lite with different specifications

Specifications of the Four Events	Median Parameter Measurements
$p(A)$=0.2, $p(B)$=0.2, $p(A \cup B)$=0.4, $p(\overline{A})$=0.8	$\mu = 0.25$, $\gamma = 0.61$
$p(A)$=0.2, $p(B)$=0.4, $p(A \cup B)$=0.6, $p(\overline{A})$=0.8	$\mu = 0.27$, $\gamma = 0.58$
$p(A)$=0.2, $p(B)$=0.6, $p(A \cup B)$=0.8, $p(\overline{A})$=0.8	$\mu = 0.25$, $\gamma = 0.60$
$p(A)$=0.4, $p(B)$=0.4, $p(A \cup B)$=0.8, $p(\overline{A})$=0.6	$\mu = 0.27$, $\gamma = 0.58$

Appendix 2.4 Classification of subjects

For each subject, we obtained four measured probability weighting functions "Prelec", "G&E", "Neo", and "Neo-Lite". For each measured function, we generated 20 decision weights for probabilities from 0.05 to 0.95, in steps of 0.05. Let $w_{FUN}(p)$ denote the sequence of generated decision weights with p=0.05, 0.10, …, 0.90, 0.95 and FUN= "Prelec", "G&E", "Neo", "Neo-Lite". We then generated another sequence $\Delta_{FUN}(p) = w_{FUN}(p) - p$. Given a weighting function FUN, a subject is classified as one of the four types below:

Inverse-S type, if $\Delta_{FUN}(0.05) > 0$ and $\Delta_{FUN}(0.95) < 0$;

S type, if $\Delta_{FUN}(0.05) < 0$ and $\Delta_{FUN}(0.95) > 0$;

Overweighting type, if $\Delta_{FUN}(p) > 0$ for all p;

Underweighting type, if $\Delta_{FUN}(p) < 0$ for all p;

Since we have confirmed that there is no subject whose measured weighting functions cross the 45-degree line (i.e. the identity line where $w(p) = p$) more than once, we think this is a proper classification of subjects based on their attitudes towards probabilities.

Appendix 2.5 Measuring probability weighting functions using binary choices in Prediction Subset A

Using the binary choice data from Prediction Subset A, where no probabilities smaller than 0.2 or larger than 0.8 are involved, we fitted the power utility function $v(x) = x^\alpha$ and three probability weighting functions: "Prelec", "G&E", "Neo". Results of parameter estimates are shown in Table 2.11.

Table 2.11 Median measurements of power utility (alpha) and probability weighting function parameters

	Prelec $w(p) = \exp(-\gamma(-\ln(p))^\mu)$	**G&E** $w(p) = \gamma p^\mu / (\gamma p^\mu + (1-p)^\mu)$	**Neo** $w(p) = \mu + \gamma p$
μ	1.16	1.46	-0.13
γ	2.00	0.50	1.12
α	1.03	0.73	0.39

Appendix 2.6 Experimental Instructions

Welcome to this experiment on decision making which will last about an hour. You will be paid £3 for coming so long as you complete all of the required tasks. In addition, you may earn up to an additional £16 depending on your decisions and

upon chance. We will pay you in cash at the end of the experiment. There are some general rules you must follow:

1. Please put away your mobile phones and do not talk to others at any time during the experiment.

2. You will use your computer to make decisions during the experiment. Do not use your mouse or keyboard to play around with the software running on your computer. If you unintentionally or intentionally close the software program running on your computer, we will ask you to leave.

3. If you have any questions during the experiment, please raise your hand. The experimenter will come to answer your questions.

Details of the Experiment

In this experiment, you will make a series of choices involving gambles and certain amounts of money. At the end of the experiment, you might play a gamble for real to determine part of your payoff from the experiment. Here is an example gamble which we refer to as Gamble A.

Gamble A

From 1 to 12	From 13 to 20
£10.00	£0.00

This gamble results in one of two possible outcomes: either winning £10 or winning nothing. The possible prizes are written in the second row of the table. The numbers in the first row of the table ("1 to 12" and "13 to 20") relate to a set of balls, numbered from 1-20 that will be used to determine the outcome of *any* gamble that you could play for real.

As you may have noticed when you enter the lab, there are 20 balls on the desk in the middle of the lab. Each ball has a number on it and there is exactly one ball with each of the numbers from 1 to 20.

At the end of the experiment, these balls will be put into a bag and one ball will be randomly drawn from this bag to determine your payoff. So for example, if you

play Gamble A at the end of the experiment, *you would get £10 if the number drawn is from 1 to 12; or you would get £0 if it is from 13 to 20.*

During the experiment, you will see gambles with different numbers of possible prizes.   For example, while gamble A has two prizes, Gamble B below has three prizes: If you were to play Gamble B for real *you would get £10 if the number drawn is 1; £7 if the number drawn is from 2 to 15; or £1 if it is from 16 to 20.*

Gamble B

From 1 to 1	From 2 to 15	From 16 to 20
£10.00	£7.00	£1.00

Finally, some options that we describe as gambles offer sure prizes.   Gamble C is an example which gives you *a sure prize of £5 whatever the number drawn from the bag.*

Gamble C

From 1 to 20
£5.00

The experiment has two parts.

Part 1: Choice Lists

In this part, you will complete some "choice lists". To understand what a choice list is, please look at Screenshot 1 which is in the 'screenshot booklet' on your desk. This shows an example choice list.

Above the example choice list there is a gamble – referred to as the "current gamble". The   example choice list is then presented as a table which, in this case, contains 11 choices each of which requires you to choose between the current gamble (which is always the Left option) or a specific amount of money (the Right option). For each choice, you should indicate your preferred option by clicking either the Left or Right circle in the "Your Choice" column.

Notice that the amounts of money given for the Right options in the choice list range, in £1 intervals, from the lowest gamble payoff (zero in this case) to the highest gamble payoff (£10 in this case). Therefore, we expect that everyone will want to choose Left in the first choice (because you might win something with Left, whereas choosing Right would guarantee that you get nothing). By contrast, for the bottom choice in this list (Choice 11) we expect that everyone will want to choose Right (because then you are guaranteed £10 whereas Left only offers a chance of that). Since the Left option stays the same moving down the choice list and the Right option gets better, we also expect that you will only switch once: in fact the computer will only allow a single switch between Left and Right in any choice list. It is, however, entirely up to you to decide where to switch from choosing Left to Right in any choice list.

Once you have completed all the choices in a list and are happy with your decisions, you should click the Confirm button to proceed. *Once you have done this you will not be able to change your confirmed choices*.

Now please look at Screenshot 2. There are two choice lists on this screenshot. Screenshot 2a on the left reproduces the example choice list that you saw in the previous screenshot but with two differences. First, it is now labelled a "Basic List". We will explain why in a moment. Secondly, notice that this basic list has been completed by a hypothetical decision maker who decided to choose Right for the first time at Choice 6. Once this decision maker had confirmed their choices for this Basic List they would then see what we call a "Zoom-in List". The relevant Zoom-in List is shown as Screenshot 2b.

Notice that via their responses to the Basic List, the decision maker has told us that they prefer the current gamble to an amount of £4 or below, but they prefer an amount of £5 or more to the current gamble. The Zoom-in list for this gamble, then presents choices between the current gamble and money amounts which range between £4 and £5 (in 10p intervals). Notice that the first and last choices of the Zoom-in list are decisions that have already been confirmed in the Basic List so these will be fixed by the computer. To complete a Zoom-in List, you must provide

decisions for the remaining 9 choices. Again, the computer will make sure that there is a single switch from Left to Right in any Zoom-in List.

In Part 1, you will be asked to make decisions about a set of different gambles using choice lists. For every gamble you will see a Basic List followed by a Zoom-in list, determined by where you switch in the Basic List.

We ask you to think about each choice in each choice list as if it were for real and as if it were the only choice you had to make. It makes sense for you to provide a considered response for each single choice because at the end of the experiment, we will randomly select one decision that you have made and use your decision to determine your payoff.

Part 2: Choices between pairs of gambles

For Part 2, you will be asked to make some choices between pairs of gambles. Now please look at Screenshot 3 which shows an example of the sort of choice you will face in Part 2. Gambles are presented in the same way as they were for Part 1 and the numbers shown in the top row of the table refer to the same set of 20 numbered balls. For each such task, you should consider which of the two gambles you would prefer and indicate your choice by clicking either the left or the right button. This will lead you to the next choice question, and you will not be able to change the choices you have made. Therefore, please think carefully before you make any click in this part. As you make your choices, please keep in mind that one of your choices will be for real.

How payment is determined

After you have finished Part 1 and 2, everyone will wait until all participants have finished their tasks. The computer will then randomly select one of your choices to be implemented for real payment.

The 'real choice' will be selected in the following way. For each participant, the computer will first randomly select either Part 1 or Part 2, with equal probabilities. If Part 2 is selected, one of the choices will be randomly selected for real payment. If

Part 1 is selected, one of the Basic Lists will be randomly selected, and one choice (i.e. one row) in that list will be randomly selected. This row gives the choice for real payment, unless it is one of your two 'switch-point' choices in that list. If the choice selected in the Basic List is one of the two 'switch-point' choices, then one of your choices in the resulting Zoom-in List will be randomly selected to be the only choice for real payment. Thus any choice you face at any point could be the one selected for real payment.

Following on from this process, the selected real choice, and what you chose in it, will be shown to you. You will receive what you chose. This may be a sure prize or it may be a gamble. In the latter case, what you receive will be determined in the following way. At the end of experiment, the experimenter will ask one of you to draw a ball from the bag of 20 numbered balls. That number will determine the payoff from *any* chosen gamble in your 'real choices'.

Before Part 1 begins, we ask you to complete a quiz. Once you have finished the quiz, click the Start Experiment button. You will only be able to start the experiment if you have answered all the quiz questions correctly. If you have any questions now or later, please raise your hand.

Chapter 3: Why People Have Non-linear Probability Weighting Functions: An Alternative Theory of Choice under Risk

3.1 Introduction

Since Prospect Theory (Kahneman & Tversky, 1979), the concept of decision weighting has been widely accepted and well-studied. Decision weights are found to be a non-linear function of probabilities, typically with overweighting of small probabilities and underweighting of moderate and large probabilities. This function is called probability weighting function and is often interpreted as reflecting an intrinsic[1] human attitude towards probabilities, driven by, for example, diminishing sensitivity towards less extreme probabilities.

However, experimental evidence that challenges this interpretation has been accumulating. For example, probability weighting has been found to depend on the domain of outcomes involved in the lotteries (i.e. whether the lotteries result in gains or losses) (see e.g. Tversky & Kahneman 1992, Abdellaoui 2000). Although theories featuring probability weighting have generally allowed for different weighting functions for gains and losses, following Cumulative Prospect Theory (Tversky & Kahneman 1992), it is not clear from these theories why probability weighting should be expected to depend on the domain of outcomes.

More recently, evidence has emerged suggesting that probability weighting is affected by the payoff sizes of lottery outcomes (see e.g. Kühberger 1998;

---

[1] 'Intrinsic' here and hereafter basically means decision-context-independent.

Etchart-Vincent 2004, 2009; Fehr-Duda et al. 2010). This means, for instance, when people are evaluating the two-outcome lottery (£*x, p*; £*y,* 1-*p*), their weighting of probability *p* would be affected by the absolute (or relative) sizes of *x* and (to) *y*. This evidence suggests that probability weighting is more than an 'intrinsic' response to chances, but also context-dependent.

Another typical contextual factor that seems to affect probability weighting is the type of task used to elicit preference (based on which we can measure probability weighting). Indirect evidence suggests that probability weighting exhibits a stable inverse-S pattern when preferences over lotteries are elicited through money valuation tasks such as Certainty-Equivalent tasks (e.g. Kahneman & Tversky 1992; Gonzalez & Wu 1999; Abdellaoui et al. 2011) but is much less regular when preferences are elicited through binary choice tasks (e.g. Hey & Orme 1994; Camerer & Ho 1994; Wu & Gonzalez 1996; Humphrey & Verschoor 2004; Conte et al. 2011). For the latter cases, other shapes of probability weighting than the inverse-S one can prevail.

In light of all these evidences discussed above, we think it would be helpful to re-think about why people are often found to have inverse-S shaped probability weighting functions and to what extent this reflects an intrinsic decisional attitude towards objective probabilities. We try to answer these questions by looking into the decision making process of choice under risk and building a model that does not pre-impose any type of probability weighting function. We propose a theory featuring an attention-based state weighting mechanism and show how our theory accommodates the evidence. This theory highlights a potentially important reason for the observation probability weighting under CPT. Briefly speaking, people weight probabilities non-linearly because their perceptions about probabilities are affected by the salience of the outcomes associated with the corresponding states. In our theory, a state is said to be the most salient if choosing differently leads to the largest welfare difference under that state.

The best example to illustrate this intuition is the case where people buy lottery tickets. When people decide to spend £1 or £2 to buy lotteries that can yield them a

million, it is not hard to imagine that they keep telling themselves "what if I win…". By doing this, people are already subconsciously overweighting the probability of winning. Their attention seems almost completely drawn to the winning state, because the winning state, with a stake of a million, is much more salient than the losing state, where the stake is only £1 or £2. Based on this idea, we formalize our theory with the following ingredients.

State-weighting (instead of probability weighting) This entails two things. First, when choices are made, lotteries are viewed as (Savage) acts instead of prospects. In other words, lotteries specify state-contingent outcomes. Second, states are evaluated one by one and the overall preference is determined by a weighted sum of the within-state comparisons. Many theories are built on the general framework of acts, such as Regret Theory (Loomes & Sugden 1982) and third-generation Prospect Theory (Schmidt et al. 2008), but not many have incorporated state-weighting (e.g. the Salience Theory of Bordalo et al. 2012).

Attention-weighted states The core ingredient of our model is that states are weighted by decision weights representing the attention of the decision maker drawn to the corresponding state. The attention paid to a state would depend on: (i) the probability of that state; (ii) how salient the outcomes associated with that state are compared to all other states. Our theory is similar to Salience Theory (Bordalo et al. 2012) in this aspect, but we have some crucial differences, which will be discussed in Section 3.2.2.

Context-dependent reference points We keep the ingredient of a reference-dependent value function from PT and CPT, but allow reference points to be context-dependent. Although we are not the first to explore this idea (e.g. Tversky & Kahneman 1991; Kőszegi & Rabin 2006; Schmidt et al. 2008), allowing the reference point to be context-dependent can greatly increase the explanatory power of our theory, given certain plausible assumptions about how reference points are context-dependent.

Though, as noted above, each of these ingredients has precedents in the literature, our combination and use of them is novel and distinctive. It provides a unified explanation for the set of findings of probability weighting (and thus for the behavioral 'anomalies' that can be explained with non-linear probability weighting) mentioned above.

To demonstrate that our theory is able to explain all these findings, we derive propositions showing that our theory is observationally equivalent to a form of CPT, the most popular theory in empirical studies of choice and probability weighting, but with specific probability-weighting functions that can depend on lottery outcomes and are tied down by our theory. Our propositions 1 and 2 show how, under our theory, the implied probability-weighting functions of CPT can depend can depend on the sizes and domains of lottery outcomes, when preferences over lotteries are elicited through the commonly used Certainty-Equivalent tasks. In Section 3.3.2 we provide intuitive analysis of how our theory need not imply the inverse-S shape of probability weighting when preference is elicited through binary choice tasks.

In general, we show that even if there is no context-independent probability weighting functions in the decision maker's cognitive system, individual behaviors can exhibit typical features that are attributed to probability-weighting in the framework of CPT. In this sense, we provide a theory of the appearance of probability weighting. However, what our analysis puts into question is not the concept of probability weighting *per se*, still less the appearance of it, but rather any interpretation under which probability weights reflect intrinsic, context-independent human attitudes towards probabilities. We think that what we call the attention-based state weighting mechanism, the key feature of our model, can be an important cause for the common observation of non-linear probability weighting.

The rest of this chapter proceeds as follows. In the next section we sketch the model. In Section 3.3 we show how our theory explains the findings about probability weighting. Section 3.4 concludes.

3.2 A Theory of Attention-Based State Weighting

3.2.1 The Model

Let preferences be defined over Savage acts. A choice problem is described by a set of states of the world S, where each state $s \in S$ occurs with objective and known probability p_s such that $\sum_{s \in S} p_s = 1$, a set of outcomes X of real numbers representing monetary payoffs, and a set of acts F, where an act $f \in F$ specifies for each state s the resulting outcome $f_s \equiv f(s) \in X$. In this chapter we consider only binary choices, i.e., a set of two acts $F = \{f, f'\}$, where at least one of the acts is a non-degenerate lottery. When an act is a degenerate lottery, or a certainty, we call it a *constant act*, whose consequences are the same in every state. The binary choice problem with *n* states can be shown in the following payoff matrix, where we choose between act f and f':

Table 3.1: The general state-payoff matrix of a decision problem

	State 1	State 2	…	State *n*
Act f	f_1	f_2	…	f_n
Act f'	f_1'	f_2'	…	f_n'

Similar to the third-generation Prospect Theory (Schmidt et al, 2008), we allow the reference point, denoted by r, to be a (degenerate or non-degenerate) lottery, so we will call r the *reference act*, with $r_s \equiv r(s) \in \mathbb{R}$. Accordingly, gains and losses will be defined state by state as $(f_s - r_s)$. The most general form of our theory consists of three functions: preference function, attention weight function, and salience function. The preference function for act f, given the alternative act f' and the reference act r, is

$$V(f, f', r) = \sum_{s \in S} \omega_s v(f_s - r_s) \tag{3.1}$$

where f_s and r_s denote respectively the monetary outcomes of act f and r in state s. The decision maker chooses f or f', depending on the relative values of $V(f, f', r)$ and $V(f', f, r)$. We use the same concept of utility (or value) function $v(.)$ as in PT and CPT. In other words, $v(.)$ is an increasing (decreasing) function of gains (losses) of monetary outcomes given a reference point, and $v(0) = 0$. Different to PT and CPT, preferences V are determined by a weighted sum of state outcomes, and the decision weights assigned to each state, ω_s, depend on the attention paid to that state, which further depends on state probabilities and state (outcome) saliences:

$$\omega_s = \frac{\varphi_s^{\eta} \cdot p_s}{\sum_{k \in S} \varphi_k^{\eta} \cdot p_k} \tag{3.2}$$

In this attention weight function, p_s is the probability of state s and φ_s is a measurement of how salient state s is, in terms of the outcomes, among all possible states. Attention paid to a state depends on the probability of the state p_s and the relative degree of salience of that state φ_s. The parameter η captures the extent to which salience distorts the decision maker's perception of probabilities. In other words, η measures the decision maker's sensitivity towards saliences. When $\eta = 0$, we have $\omega_s = p_s$ as in Expected Utility. Unless stated otherwise, we assume $\eta \geq 0$, indicating that if two states have the same probability, the more salient state gets more attention, and hence relatively a larger decision weight. It is easier to see this from the following equation, which is an immediate implication of equation (3.2) and uses States 1 and 2 for illustration.

$$\frac{\omega_1}{\omega_2} = \frac{p_1}{p_2} \cdot \left(\frac{\varphi_1}{\varphi_2}\right)^{\eta} \tag{3.3}$$

The ratio of decision weights for states s_1 and s_2 depends not only on the ratio of their probabilities, but also on their relative degree of saliences. The latter is given by

the salience function:

$$\varphi_s \equiv \varphi(f_s, f_s', r_s) = \frac{|v(f_s - r_s) - v(f_s' - r_s)|}{|v(f_{s_m} - r_s) - v(f_{s_m}' - r_s)|} \tag{3.4}$$

where $v(.)$ is the same utility function as in (3.1), f_s' denotes the payoff of act f' in state s, and s_m denotes the most salient state (i.e. the state with the largest value difference between choosing act f and choosing f'). By this definition we have $\varphi_{s_m} = 1$ and $\varphi_s < 1$ for $s \neq s_m$. A state is more salient if there is greater welfare gap between choosing one option to choosing the other. We use the distance $|v(f_s - r_s) - v(f_s' - r_s)|$ to capture this welfare gap instead of using simply the payoff difference $|f_s - f_s'|$. This is basically to allow for individual heterogeneity and the effect of reference-dependence on judging the salience of outcomes. It is also worth noticing that given a reference point r (where $r \neq f$ and $r \neq f'$), we have $\varphi(f_s, f_s') = \varphi(f_s', f_s)$ and so $\omega_s = \omega_s'$. This means that the decision weights assigned to state s is the same for both acts in the choice set. Then the decision maker would (weakly) prefer f to f' if and only if:

$$\sum_{s \in S} \omega_s \big(v(f_s - r_s) - v(f_s' - r_s)\big) \geq 0 \tag{3.5}$$

We want to again address the common ingredients and differences between our theory and PT. The latter has two key ingredients, a reference-dependent utility function and a probability weighting function, and both functions are treated as exogenous, so that a decision maker's behaviors under risk can be predicted with the knowledge of his utility function and weighting function parameters (and some assumption about the reference point). In contrast, our theory shares the same concept of an exogenous reference-dependent utility function, but drops the probability weighting function and replaces it with an attention-based state weighting mechanism (i.e. function (3.2) and (3.4)). Therefore, the only exogenous components of our preference model are the utility function $v(.)$, the reference act r, and the

parameter of salience sensitivity η. In this sense, our theory can be seen as one of 'endogenous probability weighting'.

3.2.2 Some important features of the model

Since the concepts of salience and attention-weighting are core to our model, it is worth clarifying the connections between our model and Salience Theory (Bordalo et al, 2012) in which salience is also plays a key role. Although our theory has some similarities to theirs, there is difference in crucial aspects[2]. We will use these comparisons to highlight some of the important features of our theory.

The first difference lies in how (outcome) salience is determined. Instead of proposing a specific salience function, Bordalo et al. (2012) considered three properties of the salience function: ordering, diminishing sensitivity, and reflection. The example below illustrates the definition of these properties.

Table 3.2: An example choice problem in the framework of Savage-acts

	State 1	State 2	State 3	State 4
Lottery 1	10	4	5	-5
Lottery 2	2	8	9	-9

The payoff matrix of Lottery 1 and 2 are shown in this table. According to ordering, State 1 is more salient than State 2, because the interval (4, 8) is a subset of (2, 10). According to diminishing sensitivity, State 2 is more salient than State 3, because (5, 9) = (4+1, 8+1). According to reflection, State 3 is exactly as salient as State 4, because the payoffs differ only in signs. With these properties, Salience Theory permits some predictions about ranks of state saliences just on the basis of

---

[2] Bordalo et al have a follow-up working paper (Bordalo et al. 2017) which develops salience theory in a different way to model consumer behaviors. Here we compare our model to their original Salience Theory model (Bordalo et al. 2012), as the latter is the one that ours relates to most closely.

these properties. However, this can only be done when the payoff matrix of the lotteries have specific patterns, as those in Table 3.2. For example, if the payoffs in State 3 is (5, 10) instead of (5, 9), even the three properties taken together cannot determine the relative saliences between State 2 and 3, and some assumption has to be made about the salience function.

In our theory, state saliences are determined by state payoffs, the shape of the utility function $v(.)$, and the reference point. As mentioned, we use the same concepts of utility function and reference point as in PT and CPT. Although we proposed a specific salience function (3.4), our theory is actually more general than Salience Theory, because in our theory, the three properties are implied only under certain assumptions about the utility function $v(.)$ (e.g. an increasing, concave, and symmetric utility function for gains and losses). This generalization brings the benefit of allowing for individual heterogeneity with respect to the ranking of salience, even when the payoff matrix has the pattern of Table 3.2.

The second important difference between this theory and Salience Theory lies in whether rank-dependence is imposed. In many theories of rank-dependence such as CPT, decision weights are assigned to prospect payoffs and dependent on only the probability and the rank of corresponding payoffs. This causes the well-known counter-intuition that very small changes of a payoff can lead to a dramatic change to its decision weight if the rank of that payoff among all possible payoffs also changes. In Salience Theory, decision weights are assigned to states and dependent only on the state probability and the rank of state saliences. The similar counter-intuition remains: small changes in salience can lead to dramatic changes to decision weights. These forms of rank-dependence can cause discontinuity in probability weighting and lottery valuation[3]. It is easy to see from our attention weight function and equation (3.3) that our theory avoids this problem because under equation (3.3), marginal changes in relative salience φ_s always have smooth effects on relative state weights ω_s.

Thirdly, unlike Salience Theory, our model incorporates reference-dependence,

---

[3] For a critical comment on Salience Theory (2012), see Kontek (2016).

another key ingredient in PT. More importantly, we allow reference acts to be context-dependent. This adds considerable flexibility and power to our theory in explaining the evidence of probability weighting, and provides a unified framework to account for some other behavioral anomalies such as preference reversals. Sometimes shifts of reference points can have significant influence on preferences over lotteries because as reference points change, gains and losses are coded differently. This feature is particularly relevant to our theory because it is a key tenet of behavioral models of decision making that losses 'loom larger' than gains, and shifts of reference points can generate states of losses even if all payoffs are positive. We will show later, in Section 3.3.1 and 3.3.2, that the complicated findings on probability weighting can be nicely explained by our theory with the help of some plausible assumptions about context-dependent referencing.

3.3 Implications on probability weighting

In this part we will show how the theory presented implies that the agent behaves as if according to CPT preferences, but with a probability-weighting function that is endogenous and predicted by our theory. We also show how this explains typical findings on probability weighting and some observations that are hard for CPT alone to rationalize. Since these findings have generally been the results of lab experiments, we focus on two typical types of experimental tasks used to measure probability weighting: money valuation task (e.g. CE tasks and Willingness-To-Pay tasks) and binary choice task (e.g. the method proposed by van de Kuilen & Wakker 2011).

Crucially, our theory predicts that probability-weighting will appear to take different forms in the two types of tasks. We will show that: (i) in money valuation tasks, probability weighting would be inverse-S shaped, gain-loss asymmetric, and be influenced by the level and spacing of outcomes; (ii) in binary choice tasks, probability weighting is less regular and can exhibit non-inverse-S shapes. By

implication, what pattern of probability-weighting is most commonly observed in studies using CPT will depend in part on the form of the tasks from which it is elicited.

3.3.1 Money valuations and probability weighting

Typically, to measure probability weighting functions, money valuations of lotteries, such as CE and Willingness-To-Pay (WTP for short), are elicited using valuation tasks. In such tasks, the decision maker can be seen as making a series of choices between the lottery to be valued and different sure money payoffs. In other words, sometimes it is literally true that the agent is making a series of choices between the lottery and sure sums, but, even when the task appears to have a more open-ended form, such comparisons may plausibly be taken to drive how the agent thinks about them. For example, in choice-list tasks, subjects are directly faced with multiple such choices in one list, while in WTP tasks, there might be a choice list in the decision maker's mind used to decide how much to pay for the lotteries. For this reason, we characterize money valuation tasks as a decision context where a series of choices between a lottery and a certainty are made. We denote this context as Lottery-vs-Certainty (L-vs-C for short).

What does our theory imply about choices in L-vs-C contexts? According to the intuition that greater salience attracts more attention and higher decision weight, the lottery tends to be preferred when the upside or the 'winning state' of the choice is more salient than the downside. For example, in choice between (£10, 0.1; £0, 0.9) and £1, the 'winning state' has a probability of 0.1. This 'winning state' is more salient because the former can make you better-off by £9 whereas in the 'losing state' your stake is just £1. This leads to overweighting of the probability 0.1 and a tendency to bet. By similar reasoning, the certainty tends to be preferred when the 'losing state' of the choice problem is more salient than the 'winning state', a

situation often associated with large probabilities of 'winning' (e.g. choosing between (£10, 0.9; £0, 0.1) and £9). So far, our explanation about overweighting (underweighting) of small (large) probabilities has been very simple: small (large) probabilities are overweighted (underweighted) because they are often associated with more (less) salient states in L-vs-C contexts.

Now consider moderate probabilities. People are widely found to be underweighting moderate probabilities greater than some critical value, which may be around 1/3. For instance, in choice between (£10, 0.5; £0, 0.5) and £5, people are often found to be risk averse, which is often interpreted as due to the underweighting of the probability 0.5 associated with winning £10 rather than the concavity of the utility function $v(.)$. Under our theory, this means that the 'losing state' of the choice (i.e. the state where you end up with £0 if the lottery is chosen) is often more salient than the 'winning state'. We hypothesize that this may be naturally explained by loss aversion, which can arise as a result of a shift of reference point. More generally, we think that when choosing between a lottery and a certainty, people tend to subconsciously take the certainty as a reference point, because the certainty is immediately available and tends to be seen as something already "in the pocket". For example, winning nothing while you could have got the sure £5 can feel as bad as a loss of £5 rather than as a reduced gain. Since losses 'loom larger' than gains, it is natural that the 'losing state' becomes more salient than the 'winning state' and the probability associated with the latter be underweighted.

To explore the implications of this idea and to formalize our analysis, we make the following assumption about reference points.

Assumption 1: In Lottery-vs-Certainty choice contexts, the reference act is the constant act whose consequence in every state is the certainty. For Lottery-vs-Lottery choices, the reference act is status quo, that is, the constant act whose consequence in every state is zero[4].

---

[4] There are cases where neither the certainty nor the *status quo* is likely to be taken as the reference act. For example, in Willingness-To-Accept tasks, the lottery being valued is likely to be perceived as an endowment and thus taken as the reference act. Our analysis here focuses on situations where the lottery is not perceived as an endowment.

Without loss of generality, let c be the money valuation[5] of a lottery $(x, p; y, 1 - p)$, where $x > c > y \geq 0$ and $1 > p > 0$, so that the decision maker is indifferent between receiving c and playing the lottery (see Table 3.3).

Table 3.3: The money valuation decision problem (positive payoffs)

| | State 1 | State 2 |
	p	1-p
The Lottery	x	y
Certainty (reference point)	c	c

According to our theory, this indifference implies the equation

$$\omega_1 v(x - c) + \omega_2 v(y - c) = 0 \tag{3.6}$$

For the convenience of deriving theoretical results, we make a standard assumption about the Prospect-Theory-type utility function with loss aversion.

Assumption 2:

$$v(x) = \begin{cases} (x - r)^\alpha \ , & x \geq 0 \\ -\lambda(-(x - r))^\alpha \ , & x < 0 \end{cases}$$

where r is the consequence of the reference act, $\lambda > 0$ is the loss aversion parameter, and $\alpha > 0$ is the power parameter capturing the curvature of the utility function. We further assume that the power parameter of the utility function is the same for gains and for losses. This assumption greatly simplifies our analysis.

To begin our exposition of how our theory accounts for the evidence about probability weighting, we present Proposition 1. This proposition shows that, when we consider positive payoffs and money valuations of lotteries, our theory is observationally-equivalent to a particular parameterized form of CPT. We use CPT

---

[5] When we say 'money valuation' of a lottery, we generally refer to the Certainty-Equivalents of the lottery.

basically because it has been one of the most widely used theories in empirical studies of probability weighting. The key message from this proposition is that, if a decision maker behaves according to our theory but is treated as a CPT agent with his preferences elicited through money valuation tasks, she will be found to have the weighting function *w* (for proofs see Appendix 3.1).

Proposition 1 *Consider an agent's valuation of the lottery* $(x, p; y, 1 - p)$, *where* $x > y \geq 0$ *and* $1 > p > 0$. *The following two statements are equivalent:*

(i) The agent has a utility function $v(x)$ *as specified in Assumption 2 and his valuation of the lottery is determined by equations (3.2), (3.4), (3.6);*

(ii) The agent has the same utility function $v(x)$ *but behaves according to Cumulative Prospect Theory with* $r = 0$ *(i.e. status quo) and with the weighting function:*

$$w(p) = \frac{[g(p)x + (1 - g(p))y]^\alpha - y^\alpha}{x^\alpha - y^\alpha}$$

where

$$g(p) = \frac{p^{\frac{1}{\alpha(\eta+1)}}}{\lambda^{\frac{1}{\alpha}}(1 - p)^{\frac{1}{\alpha(\eta+1)}} + p^{\frac{1}{\alpha(\eta+1)}}}$$

For clarity, we will refer to the weighting function, *w(p)*, in part (ii) of the Proposition as "the implied probability weighting function". Proposition 1 shows that the implied probability weighting function also depends on non-probabilistic factors: lottery payoffs *x, y*, utility function parameter[6] α, degree of loss aversion λ, and degree of salience sensitivity η. In this sense, the implied probability weighting function is endogenous under our theory. Moreover, it reflects payoffs in a different

---

[6] It may need to be clarified that, theoretically, the value of α in our theory is the same as the α in Cumulative Prospect Theory. This is natural because both share the same concept of utility function. Crucially, this concept implies that the utility function captures individuals' exogenous attitudes towards outcomes. Based on this exogenous part, CPT introduces another independent part (the weighting function) to model preferences, whereas we introduce a component of decision making process that can be seen as the implied probability weighting function $w(p)$. In neither case is the shape of the utility function affected by the shape of the weighting function.

way from the rank-dependent structure of CPT and, rather than being a primitive intrinsic response to probability, it reflects other preference parameters, including the salience sensitivity parameter which is distinctive to our theory and does not feature in standard formulations of CPT. The function $g(p)$ resembles the two-parameter probability weighting function proposed by Goldstein and Einhorn (1987), which exhibits the typical inverse-S shape with $x > y \geq 0$ and with a wide range of parameter values. Therefore, as we will elaborate below, the function $w(p)$ can also exhibit inverse-S shapes with typical parameter values.

Figure 3.1 below presents the implied probability weighting function $w(p)$, for different values of the parameters of our preference model and for different values of outcomes x and y. The first panel holds all terms constant at particular values, except for λ which is varied. Similarly, second panel varies η only, the third varies α only, and the fourth varies only the payoff ratio x/y. Specifically, Figure 3.1a illustrates that λ seems to influence only the elevation of the weighting function, with greater loss aversion associated with less elevation (or more pessimism). Figure 3.1b illustrates that the salience sensitivity parameter η seems to affect the degree of curvature of the weighting function (or likelihood insensitivity), with higher η related to greater likelihood insensitivity. Figure 3.1c illustrates that utility function parameter α seems to have an impact on both elevation and curvature of the weighting function, with w seemingly more elevated and less curved when the utility function for gains is more concave.

Although it is easy to see that the implied probability weighting function now depends on the outcomes x and y, it is only affected by the relative outcome spacing, or the ratio x/y. For example, doubling both x and y would have no impact on w. As illustrated in Figure 3.1d, the larger the ratio x/y is, the more elevated the weighting is, indicating greater optimism about the upside of risk. It is worth mentioning that the weighting function of Proposition 1 has a specific implication about how relative outcome spacing x/y affects the decision weighting: For individuals with concave utility functions (i.e. $0 < \alpha < 1$), their weighting functions are generally more elevated
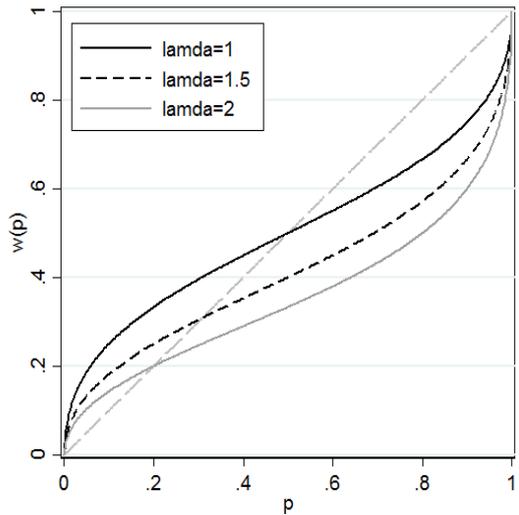
$\alpha = 1, \; \eta = 1 \;\; y = 0$

Figure 3.1a: Manipulating λ
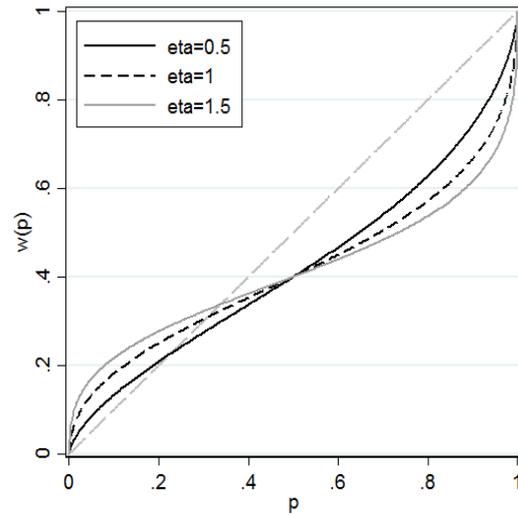
$\alpha = 1, \; \lambda = 1.5, \; y = 0$

Figure 3.1b: Manipulating η
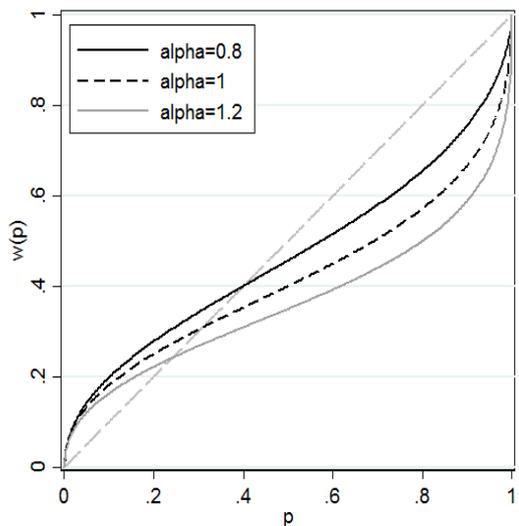
$\eta = 1, \; \lambda = 1.5 \;\; y = 0$

Figure 3.1c: Manipulating α
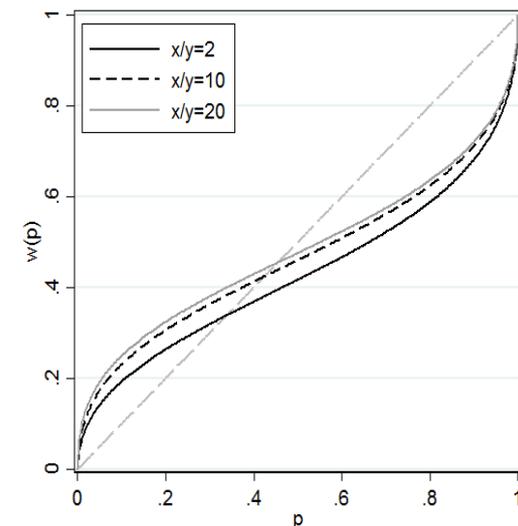
$\alpha = 0.8, \; \eta = 1, \; \lambda = 1.5, \; y \neq 0$

Figure 3.1d: Manipulating x/y

Figure 3.1: Plotting $w(p)$ and manipulate different parameter values

for (two-outcome) lotteries with larger relative outcome spacing, whereas for those with convex utility functions (i.e. $\alpha > 1$), their weighting functions are generally less elevated for (two-outcome) lotteries with larger relative outcome spacing[7].

---

[7] Although this implication is perhaps true only in our model and only when the utility function has the simplest power functional form, it points to an intuitive idea that the way probability weighting is sensitive to lottery

In general Proposition 1 gives an account of why the commonly-found inverse-S probability weighting function might have that form and why and how probability weighting are found to be affected by the spacing or relative sizes of outcomes. It is also implied that individual heterogeneity of the weighting function found in experiments may be closely related to individual difference in their utility function curvature α, degree of salience sensitivity η, and degree of loss aversion λ.

So far we have been considering positive lottery outcomes. Next consider negative outcomes[8]. Let c^- denote the decision maker's money valuation of the lottery $(x, p; y, 1-p)$, where $x < y \leq 0$ and $1 > p > 0$.

Table 3.4: The money valuation decision problem (negative payoffs)

	State 1 p	State 2 1-p
The Lottery	x	y
Certainty (reference point)	c^-	c^-

Under our theory, this indifference means the equation

$$\omega_1 v(x - c^-) + \omega_2 v(y - c^-) = 0 \tag{3.7}$$

Our result regarding negative payoffs and probability weighting is shown in the proposition below (proofs provided in Appendix 3.2), which also gives a CPT-equivalent of our theory in the domain of loss. It shows how our theory can explain a specific pattern of difference between probability weighting for gains and losses, under Assumption 1 and 2.

---

outcomes would depend on the curvature of the decision maker's utility function over the outcomes or attitude towards the outcomes. Proof of this implication is provided in Appendix 3.3.

[8] Here positive outcomes basically refer to the case where lottery outcomes are gains of money, and negative outcomes refer to the case where lottery outcomes are losses of money. But, under our theoretical framework, *in each case* it is possible for an outcome to lie either side of the reference-point. For the decision maker, a positive outcome can be treated as a loss if it lies below her reference point, and a negative outcome can be treated as a gain if it lies above the reference point.

Proposition 2 *Consider an agent's valuation of the lottery* $(x, p; y, 1-p)$*, where* $x < y \le 0$ *and* $1 > p > 0$*. The following two statements are equivalent:*

(i) The agent has a utility function $v(x)$ *as specified in Assumption 2 and his valuation of the lottery is determined by equations (3.2), (3.4), (3.7);*

(ii) The agent has the same utility function $v(x)$ *but behaves according to Cumulative Prospect Theory with* $r = 0$ *(i.e. status quo) and with the probability weighting function:*

$$w(p)^- = \frac{[-g(p)^- x - (1 - g(p)^-)y]^\alpha - (-y)^\alpha}{(-x)^\alpha - (-y)^\alpha}$$

where

$$g(p)^- = \frac{p^{\frac{1}{\alpha(\eta+1)}}}{\lambda^{-\frac{1}{\alpha}}(1-p)^{\frac{1}{\alpha(\eta+1)}} + p^{\frac{1}{\alpha(\eta+1)}}}$$

We can see by comparing Proposition 2 to Proposition 1 that the difference between the implied probability weighting for gains and losses comes from loss aversion[9]. Loss aversion is relevant both when *x, y > 0* and when *x, y < 0*, because, in both cases, a consequence may exceed or fall short of the reference point. But, as Propositions 1 and 2 show, how λ enters the implied probability weighting function is subtly different in the two cases.

When there is no loss aversion (i.e. $\lambda = 1$), $g(p) = g(p)^-$, and the shape of $w(p)$ and $w(p)^-$ will be exactly the same given the values of the other parameters. When $\lambda > 1$, $\lambda^{-\frac{1}{\alpha}} < \lambda^{\frac{1}{\alpha}}$ for any $p \in (0,1)$ and any $\alpha > 0$, so that $g(p)^- > g(p)$ for any given p. Since $w(.)$ is an increasing function of $g(p)$, $g(p)^- > g(p)$ entails $w^- > w$, for any p. In other words, under Assumption 1, our theory predicts that probability weighting for losses is generally more elevated (implying greater pessimism) than for gains, as illustrated in Figure 3.2. This is very intuitive because under CPT, the probability weighting function depicts probability weights

---

[9] Note the difference between the denominators of $g(p)$ and $g(p)^-$. The two weighting functions $w(p)$ and $w(p)^-$ are identical, except for having different $g(.)$ functions. And that the only difference in the latter is the sign of the exponent on λ in the denominator.

associated with the largest gain or the largest loss, which corresponds to State 1 in
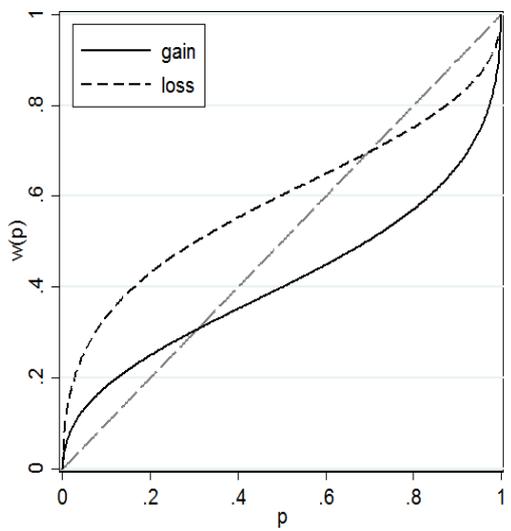
$\alpha=0.8, \ \eta=1, \ \lambda=1.5, \ y=0$

Figure 3.2a
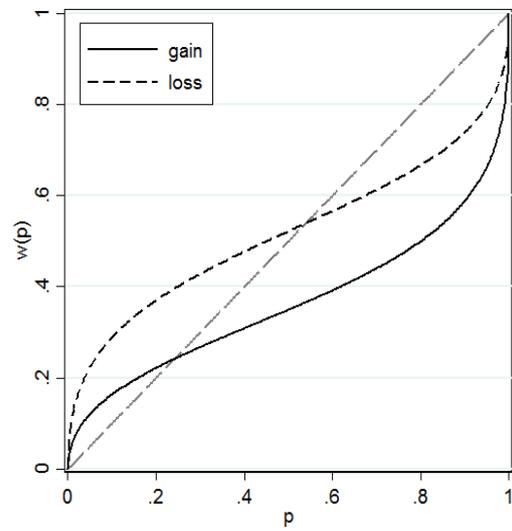
$\alpha=1.2, \ \eta=1, \ \lambda=1.5, \ y=0$

Figure 3.2b
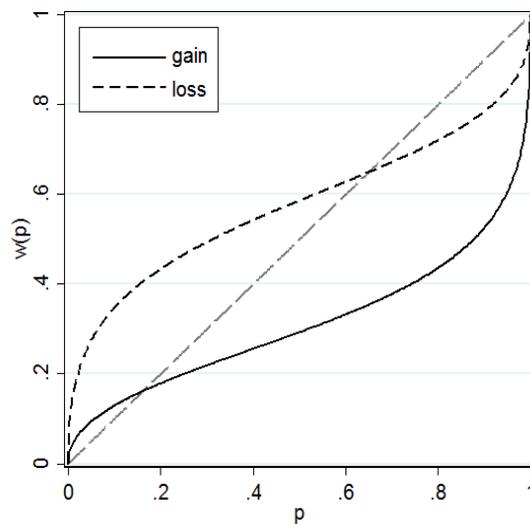
$\alpha=1.2, \ \eta=1, \ \lambda=2, \ y=0$

Figure 3.2c

Figure 3.2: Comparing $w(p)$ and $w(p)^-$ with specific parameter values

Table 3.3 and Table 3.4. However, State 1 is a state of gain in the case of gain (Table 3.3) but a state of loss in the case of loss (Table 3.4). Due to loss aversion, the State 1 in Table 3.3 is less salient than the State 1 in Table 3.4 given any probability p. This

implies a more elevated weighting function for loss under CPT.

Figure 3.2 also illustrates that the size of the gap seems to be larger either when α is smaller (to see this compare Figure 3.2a and 3.2b) or when λ is larger (to see this compare Figure 3.2b and 3.2c). In the literature, while probability weighting has been widely found to be inverse-S shaped for both gains and losses, the difference between these two seems to be much less patterned. Although there are many findings that the weighting function for loss is more elevated than that for gains (e.g. Tversky & Kahneman 1992, Abdellaoui 2000, Kusev et al 2009) as implied by Proposition 2, there are also evidence where it is not (e.g. Lattimore et al 1992, Bruhin et al 2010).

We speculate that this mixture of evidence might be accounted for once we relax or change our assumptions underlying Proposition 2. For example, besides loss aversion, another source of the gap between probability weighting for gains and losses may be the asymmetry in the utility function curvature (i.e. allowing α to be different for gains and losses), which has been typical in the literature. In addition, if the domain of outcomes turns out to be another contextual factor influencing the determination of reference points (i.e. allowing Assumption 1 to be different for gains and losses), there can also be more variations of this gap, depending on how the domain of outcomes affect the determination of reference points.

Unlike other theories that simply allow for probability weighting to be different for gains and losses, our theory provides a framework within which such a difference is endogenous and predicted. Proposition 1 and 2 together show how one of those factors, loss aversion, can explain one of the most prevalent patterns found in the literature that probability weighting for loss is generally more elevated than for gains. We have shown that a typical decision maker, who has a slightly concave utility function (e.g. α around 0.8), is slightly loss-averse (e.g. λ around 1.5), and is moderately sensitive to payoff salience (e.g. η around 1), can exhibit the typical patterns of empirically observed probability weighting functions.

3.3.2 Binary choices and probability weighting

As has been shown, our theory implies the appearance of inverse-S shaped probability weighting in the case of money valuations because there small (large) probabilities are often associated with more (less) salient states when a two-outcome lottery is being evaluated against some sure amounts of money. An interesting follow-up question would be, how about other cases than money valuations? In this part we discuss our theoretical implications about probability weighting in the case of binary choices between lotteries, or specifically the case where probability weighting is measured using binary choice data. We denote this decision context as Lottery-vs-Lottery (L-vs-L for short). Although money valuation data such as Certainty-Equivalents is popular for researchers, a significant part of empirical studies of probability weighting use binary choice data (e.g. Lattimore et al. 1992, Harrison & Rutström 2009, Hey et al. 2010). Somewhat surprisingly, in most of these studies, the inverse-S pattern of probability weighting is not dominant.

Our theory implies distinctive features of probability weighting in the case of binary choices. Specifically, probability weighting can be much less regular and the weighting function can exhibit a variety of shapes depending on the lotteries used in the tasks. This result is based on two differences between the L-vs-C context (money valuations) and the L-vs-L context (binary choices). First, the *status quo* seems to be a more natural reference point when there is no certainty in the choice set (as specified in Assumption 1). Second, as the number of state increases from two to more, in the latter case, states with small probabilities need no longer be associated with greater salience. For example, consider two stochastically independent lotteries A= ($x_1, p; \ x_3, 1-p$) and B= ($x_2, q; \ x_3, 1-q$) where $x_1 > x_2 > x_3$ and $p < q$. As will be clear in Table 3.5 below, probability p is associated with two states 1 and 2. Even if p is small and State 2 is salient enough to draw disproportionally high attention (i.e. to be overweighted), State 1 may be underweighted because it can be less salient than State 2 and State 3. As a result, the net effect of outcome salience on probability p

may be positive or negative, leading to respectively overweighting or underweighting. Following this reasoning, small probabilities are no longer always overweighted and large probabilities need not be underweighted.

Table 3.5:   A Lottery-vs-Lottery choice problem

	State 1 pq	State 2 $p(1-q)$	State 3 $(1-p)q$	State 4 $(1-p)(1-q)$
Lottery A	x_1	x_1	x_3	x_3
Lottery B	x_2	x_3	x_2	x_3

Therefore, we can show that, under CPT, a decision maker can be found to have a non-inverse-S probability weighting function in this case[10]. Specifically, the shape of the underlying weighting function will depend on the outcomes and probabilities of the lotteries in the choice set. Take Lottery A and B in Table 3.5 as an example. Our theory predicts that, ceteris paribus, the probability p is more likely to be overweighted (underweighted) when x_2 is closer (further) to x_3. This is because when x_2 is closer to x_3, the difference between x_1 and x_2 is larger and State 1 becomes more salient. The probability p is also more likely to be overweighted (underweighted) when, ceteris paribus, q is smaller (larger). This is because State 1 occurs with probability pq. Suppose State 1 is assigned a distorted decision weight, i.e. a decision weight higher than its probability pq, due to its outcome salience. The larger q is relative to p, the more this distortion can be seen as attributed to the distorted probability weight of q than to the distorted probability weight of p. Symmetrically, whether probability q is overweighted or underweighted depends on all the other lottery parameters, (i.e, x_1, x_2, x_3, and p) in a similar way.

An interesting implication of this result is that we can potentially manipulate the shape of the underlying probability weighting function we try to measure by manipulating the set of lottery parameters (x_1, x_2, x_3, p, q) we use to elicit preferences. Moreover, we can roughly identify conditions of these parameters for

---

[10] For analytical results see Appendix 3.4.

different patterns of probability weighting. For example, we would expect that, fixing x_1 and x_3, there exists a set of (x_2, p, q) such that probability p is underweighted and probability q is overweighted, which is consistent with the S shape of probability weighting[11]. The sensitivity of probability weighting to the lottery parameters also means that preferences between lotteries are sensitive to these parameters. For example, for a typical choice of trade-off where people choose between a small probability to win a large prize and a large probability to a small prize, we predict that preferences can be extremely sensitive to the relative sizes of the prizes and probabilities. It may be a mistake to think either that most people would choose the latter because of risk aversion or that they would prefer most of the time the former because small probabilities are overweighted.

3.4 Concluding Remarks

3.4.1 A summary

In this chapter we propose a potential explanation for why, under CPT, people appear to have non-linearly probability weighting functions and why these functions are typically found to be: (i) inverse-S shaped, (ii) different for the domain of gains and losses, (iii) sensitive to the spacing or relative sizes of outcomes. There is also evidence, though less well-known, that the patterns of measured probability weighting seem to be sensitive to the type of tasks used to elicit preferences. Some of these evidences cannot be well accounted for by PT or CPT, or any theory that imposes a single context-independent probability weighting function.

We provide a theory of attention-based state weighting as our explanation and have shown how this theory fits into the findings about probability weighting. Rather than proposing specific functional forms for probability weighting, with Proposition

---

[11] For analytical results see Appendix 3.4.

1 and 2 we aim at illustrating, under some specific assumptions, that what has been seen as probability weighting (and also the inverse-S shape of the probability weighting function) may simply be a result of the attention-based state weighting mechanism, in which deviations of decision weights from probabilities come from disproportionate allocations of decisional attention.

The two propositions show that an agent who behaves according to our theory can exhibit the typical inverse-S probability weighting function with its shape determined by the agent's utility function parameter, degree of loss aversion, salience sensitivity, the reference point, and the level and spacing of outcomes, when probability weighting is measured from money valuation tasks. In addition, Section 3.3.2 presents an intuition for the implication that the implied probability weighting function can exhibit any shape depending on the agent's utility function parameter, salience sensitivity, and the lottery probabilities and payoffs, when probability weighting is measured using binary choice tasks. The key message from that section are that probability weighting can be sensitive to decision contexts (such as to the types of experimental tasks) and that our theory can provide an account for such context-sensitivity.

Again, our theoretical results challenge the popular understanding of decision weighting as reflecting a certain type of intrinsic attitudes towards probabilities. We prefer to interpret decision weighting as a result of both intrinsic imprecise perception of probabilities and the attention-based state-weighting heuristic induced in typical decision contexts under risk. The former makes it possible for the latter to play a role. We think that in light of the accumulating experimental findings of context-dependent decision weighting, our theory provides a more intuitive and systematic account for the evidence than a theory that simply imposes different weighting function for different decision contexts.

3.4.2 Prospects for experimental tests

The theory we present has a wide scope for experimental tests. The first testable implication is that decision weights are affected by not only outcome ranks but also (relative) outcome spacing, as implied by Proposition 1. Most straightforwardly, since our theory has a CPT-equivalent (Proposition 1 and 2), it can explain most of the experimental evidences which CPT could explain, and also some of the evidence that CPT couldn't. This is mainly because our theory implies the appearance of probability weighting in forms that vary with factors that would be irrelevant under CPT. Therefore we consider our theory a 'best-buy' alternative of CPT. We think that our theory can do better, both descriptively and predictively, in fitting experimental data of choice under risk, with an equal or even less number of parameters than what CPT would require.

It would be more interesting to test the novel implications of our theory. One important implication is the differentiation between valuation tasks and choice tasks in terms of their probability weighting patterns. We suggest that the inverse-S pattern of probability weighting is rather stable for money valuation tasks, but very sensitive to the lottery parameters in choice tasks. Consequently, we can vary the lotteries used in choice tasks to manipulate the probability weighting function measurements we obtain from the choice tasks. We think this is a simple way to test our theory.

Another interesting scope for test relates to the well-known preference reversal phenomenon. Our theory provides an explanation for the standard pattern of preference reversal: its persistence is largely due to a stable and dominant preference for the $-bet (i.e. a small probability to win a large prize) over the P-bet (a large probability to win a small prize) in money valuation tasks. This preference is stable and dominant because when preferences are revealed through money valuation tasks, small (large) probabilities tend to be seen as overweighted (underweighted). Therefore, people always have a tendency to put a higher money value on the $-bet. However, our theory predicts that preferences (for the P-bet) in choice tasks are less dominant and can easily be altered when the lottery parameters are different. As a result, the rate of standard preference reversal can somehow be manipulated by

manipulating the lotteries used to elicit preferences. Furthermore, our theory predicts that the preference for the $-bet may become much less dominant if we use non-money valuation tasks. For example, if we used probabilistic valuations instead, where people evaluate a lottery with another lottery, the decision context essentially becomes a Lottery-vs-Lottery case and the prevalence of the standard preference reversal may disappear[12].

Finally, we think that our assumption about context-dependent reference points also generates some interesting implications. For instance, under Assumption 1, we would find a potential link between loss aversion and probability weighting, particularly the component representing the degree of pessimism or optimism. However, our theory predicts that loss aversion plays a role only in money valuation tasks but not in binary choice tasks, because of a shift of reference points (from the *status quo*) to the sure money options. Although it is certainly unsurprising that a change of reference points often leads to systematic changes of behaviors, we think it is important to recognize that in experimental settings subjects' reference points often deviate from the *status quo* and that these deviations often follow some regularities. It would be very helpful to identify these regularities and utilize them to explain experimental behaviors.

---

[12] Cubitt et al (2004) use probabilistic valuation tasks as well as monetary valuation tasks, and found markedly different patterns of preference reversal relative to choice in the two cases. See also Butler and Loomes (2007) for a similar result.

Appendices for Chapter 3

Appendix 3.1 Proof of Proposition 1

For Proposition 1, let c denote the decision maker's money value of the lottery $(x, p; y, 1-p)$, where $x > y \geq 0$ and $1 > p > 0$. So probability weighting is measured based on indifference between the lottery and c. The following table is the same as Table 3.3.

	State 1 p	State 2 $1\text{-}p$
The Lottery	x	y
Certainty (reference point)	c	c

We first start with our theory. The indifference implies the equation

$$\omega_1 v(x - c) + \omega_2 v(y - c) = 0$$

which is, under Assumption 1 and 2,

$$\omega_1 v(x - c) - \omega_2 \lambda v(c - y) = 0$$

$$or$$

$$\frac{\omega_1}{\omega_2} = \frac{\lambda v(c - y)}{v(x - c)} \tag{3.8}$$

According to equation (3.3), the left-hand side of equation (3.8) is equal to

$$\frac{p}{1 - p} \cdot \left(\frac{\varphi_1}{\varphi_2}\right)^{\eta}$$

where $\frac{\varphi_1}{\varphi_2} = \frac{v(x-c)}{\lambda v(c-y)}$. Equation (3.8) can be written as

$$\frac{p}{1 - p} \cdot \left(\frac{v(x - c)}{\lambda v(c - y)}\right)^{\eta} = \frac{\lambda v(c - y)}{v(x - c)}$$

Rearrange this we have

$$\frac{p}{1 - p} = \left(\frac{\lambda v(c - y)}{v(x - c)}\right)^{\eta+1} \tag{3.9}$$

Under our theory, equation (3.9) can be seen as the simplified preference condition

for the decision problem in Table 3.4. Notice that equation (3.9) holds regardless of whether State 1 or State 2 is more salient to the decision maker. Now under the assumption of the power utility function $v(.)$, we have

$$\frac{p}{1-p} = \left(\frac{\lambda(c-y)^\alpha}{(x-c)^\alpha}\right)^{\eta+1}$$

Now we do a series of rearrangements:

$$\frac{p}{1-p} = \left(\frac{\lambda(c-y)^\alpha}{(x-c)^\alpha}\right)^{\eta+1}$$

$$\Rightarrow \quad \left(\frac{p}{1-p}\right)^{\frac{1}{\alpha(\eta+1)}} = \frac{\lambda^{\frac{1}{\alpha}}(c-y)}{x-c}$$

$$\Rightarrow \quad \left(\frac{p}{1-p}\right)^{\frac{1}{\alpha(\eta+1)}}(x-c) = \lambda^{\frac{1}{\alpha}}(c-y)$$

$$\Rightarrow \quad c = \frac{\left(\frac{p}{1-p}\right)^{\frac{1}{\alpha(\eta+1)}}x + \lambda^{\frac{1}{\alpha}}y}{\left(\frac{p}{1-p}\right)^{\frac{1}{\alpha(\eta+1)}} + \lambda^{\frac{1}{\alpha}}}$$

which is further equivalent to

$$c = \frac{p^{\frac{1}{\alpha(\eta+1)}}}{p^{\frac{1}{\alpha(\eta+1)}} + \lambda^{\frac{1}{\alpha}}(1-p)^{\frac{1}{\alpha(\eta+1)}}}x + \frac{\lambda^{\frac{1}{\alpha}}(1-p)^{\frac{1}{\alpha(\eta+1)}}}{p^{\frac{1}{\alpha(\eta+1)}} + \lambda^{\frac{1}{\alpha}}(1-p)^{\frac{1}{\alpha(\eta+1)}}}y$$

Now define

$$g(p) \equiv \frac{p^{\frac{1}{\alpha(\eta+1)}}}{\lambda^{\frac{1}{\alpha}}(1-p)^{\frac{1}{\alpha(\eta+1)}} + p^{\frac{1}{\alpha(\eta+1)}}} \tag{3.10}$$

Since we have assumed that $\lambda > 0$, $\alpha > 0$, $\eta \geq 0$, and $0 < p < 1$, we would have $0 < g(p) < 1$. Now equation (3.9) is equivalent to:

$$c = g(p)x + \left(1 - g(p)\right)y \tag{3.11}$$

Take the power of α to both sides of equation (3.11) we have

$$c^\alpha = [g(p)x + \left(1 - g(p)\right)y]^\alpha \tag{3.12}$$

Now let the right-hand side of (3.12) be equal to $w(p)x^\alpha + \left(1 - w(p)\right)y^\alpha$, where

$w(p)$ is a probability weighting function as in CPT. We can then solve for w:

$$w = \frac{[g(p)x + (1 - g(p))y]^\alpha - y^\alpha}{x^\alpha - y^\alpha} \tag{3.13}$$

In other words, equation (3.12) is equivalent to

$$v(c) = w(p)v(x) + (1 - w(p))v(y) \tag{3.14}$$

where $v(.) = (.)^\alpha$ and $w(p)$ specified as in equation (3.13).

Now we start with CPT, i.e. equation (3.14), (3.13) and (3.10), and show that the preference condition of our theory, i.e. equation (3.9), can be obtained.

With power utility function and equation (3.12), (3.13) can be written as

$$c^\alpha = \frac{[g(p)x + (1 - g(p))y]^\alpha - y^\alpha}{x^\alpha - y^\alpha} x^\alpha + \frac{x^\alpha - [g(p)x + (1 - g(p))y]^\alpha}{x^\alpha - y^\alpha} y^\alpha$$

which can be simplified to

$$c^\alpha(x^\alpha - y^\alpha) = [g(p)x + (1 - g(p))y]^\alpha(x^\alpha - y^\alpha)$$

Since $x > y$ and $\alpha > 0$, we have $x^\alpha - y^\alpha > 0$. Hence

$$c^\alpha = [g(p)x + (1 - g(p))y]^\alpha$$

And since $c > 0$ and $g(p)x + (1 - g(p))y > 0$, we can remove the power on both sides to get equation (3.11):

$$c = g(p)x + (1 - g(p))y \tag{3.11}$$

Since $0 < p < 1$, $g(p)$ can be written as, by dividing both the numerator and denominator by $(1 - p)^{\frac{1}{\alpha(\eta+1)}}$,

$$g(p) = \frac{\left(\dfrac{p}{1 - p}\right)^{\frac{1}{\alpha(\eta+1)}}}{\left(\dfrac{p}{1 - p}\right)^{\frac{1}{\alpha(\eta+1)}} + \lambda^{\frac{1}{\alpha}}} \tag{3.15}$$

Combining equation (3.11) and (3.15), we have

$$c = \frac{\left(\frac{p}{1-p}\right)^{\frac{1}{\alpha(\eta+1)}} x + \lambda^{\frac{1}{\alpha}}y}{\left(\frac{p}{1-p}\right)^{\frac{1}{\alpha(\eta+1)}} + \lambda^{\frac{1}{\alpha}}}$$

Rearrange this we have

$$\left(\frac{p}{1-p}\right)^{\frac{1}{\alpha(\eta+1)}} (x - c) = \lambda^{\frac{1}{\alpha}}(c - y)$$

Since x>c, this is equivalent to

$$\left(\frac{p}{1-p}\right)^{\frac{1}{\alpha(\eta+1)}} = \frac{\lambda^{\frac{1}{\alpha}}(c - y)}{x - c}$$

Take the power of $\alpha(\eta + 1)$ to both sides and we have equation (3.9)

$$\frac{p}{1-p} = \left(\frac{\lambda v(c - y)}{v(x - c)}\right)^{\eta+1}$$

By this we have shown that whether we start with CPT or our theory, we end up with the other preference condition. So Proposition 1 is proved. ∎

Appendix 3.2 Proof of Proposition 2

Now consider the domain of loss. let c^- denote the decision maker's money value of the lottery $(x, p; y, 1 - p)$, where $x < y \le 0$ and $1 > p > 0$. The following table is the same as Table 3.4

	State 1 p	State 2 $1\text{-}p$
The Lottery	x	y
Certainty (reference point)	c^-	c^-

Under our theory, this indifference means the equation

$$\omega_1 v(x - c^-) + \omega_2 v(y - c^-) = 0$$

which is now equivalent to

$$-\omega_1 \lambda v(x - c^-) + \omega_2 v(c^- - y) = 0 \quad or \quad \frac{\omega_1}{\omega_2} = \frac{v(c^- - y)}{\lambda v(x - c^-)}$$

Following similar procedures as we do in the proof of Proposition 1, we have the rearranged preference condition in this case:

$$c^- = g(p)^- x + (1 - g(p)^-) y \tag{3.16}$$

where

$$g(p)^- = \frac{p^{\frac{1}{\alpha(\eta+1)}}}{\lambda^{-\frac{1}{\alpha}}(1-p)^{\frac{1}{\alpha(\eta+1)}} + p^{\frac{1}{\alpha(\eta+1)}}} \tag{3.17}$$

Again, consider the CPT preference condition in this case. There is a weighting function for loss, $w(p)^-$, and a value function as specified in Assumption 2, such that

$$\lambda(-c^-)^\alpha = w(p)^- \lambda(-x)^\alpha + (1 - w(p)^-) \lambda(-y)^\alpha \tag{3.18}$$

Further rearrange (3.16) to be

$$(-c^-)^\alpha = [g(p)^-(-x) + (1 - g(p)^-)(-y)]^{\ \alpha} \tag{3.19}$$

Let equations (3.18) and (3.19) be equivalent and solve for $w(p)$, we have

$$w(p)^- = \frac{[-g(p)^- x - (1 - g(p)^-) y]^\alpha - (-y)^\alpha}{(-x)^\alpha - (-y)^\alpha} \tag{3.20}$$

Hence, starting with our theory, we have obtained an equivalent representation of CPT, i.e. equations (3.18) and (3.20). Similar to our proof of Proposition 1, we can also start with equations (3.17), (3.18), and (3.20) and obtain the preference condition of our theory. Proposition 2 can be proved accordingly.

Appendix 3.3 Proof of Proposition 1's implication

Proposition 1 implies that, for two-outcome lotteries (*x, p; y,* 1-*p*), where *x>y>*0 and 0<*p*<1, the decision weight of *x* (i.e. function *w*), is an increasing function of the ratio *x/y* when $0 < \alpha < 1$ (i.e. concave utility function), but a decreasing function of the ratio *x/y* when $\alpha > 1$. Here's the proof.

Let $x/y = z$. Then w can be rewritten as

$$w = \frac{[g(p)z + 1 - g(p)]^\alpha - 1}{z^\alpha - 1}$$

For a given probability p and a given decision maker, we can treat $g(p)$ and α as exogenous parameters, so that $w(.)$ can be seen as a function of z:

$$w(z) = \frac{(gz + 1 - g)^\alpha - 1}{z^\alpha - 1}$$

Since $g \in (0,1)$, $\alpha > 0$, and $z > 1$, $w(z)$ is differentiable. Then we check the FOC for its monotonicity.

$$w'(z) = \frac{\alpha g(z^\alpha - 1)(gz + 1 - g)^{\alpha-1} - \alpha z^{\alpha-1}((gz + 1 - g)^\alpha - 1)}{(z^\alpha - 1)^2}$$

Since $(z^\alpha - 1)^2 > 0$, we have $w'(z) \lessgtr 0$ if and only if

$$\alpha g(z^\alpha - 1)(gz + 1 - g)^{\alpha-1} \lessgtr \alpha z^{\alpha-1}((gz + 1 - g)^\alpha - 1)$$

Now let $h = gz + 1 - g$. Since $z > 1$ and $g \in (0,1)$, we have $z > h > 1$. The inequality above now becomes

$$\alpha g(z^\alpha - 1)h^{\alpha-1} \lessgtr \alpha z^{\alpha-1}(h^\alpha - 1)$$

Rearrange this we have

$$g \lessgtr \frac{h - \dfrac{1}{h^{\alpha-1}}}{z - \dfrac{1}{z^{\alpha-1}}}$$

Now define function $f(t) = t - \frac{1}{t^{\alpha-1}}$ with the domain $t \geq 1$. Obviously this function is continuous and at least two times differentiable on its domain. Now we check its concavity/convexity.

$$f''(t) = \alpha(1 - \alpha)t^{-\alpha-1}$$

Since $\alpha > 0$ and $t \geq 1$, we know that $t^{-\alpha-1} > 0$, so that the sign of $f''(t)$ depends on the sign of $(1 - \alpha)$.

When $\alpha > 1$, $f''(t) < 0$, function $f(t)$ is concave on the domain. According to concavity, we have that for any t_1, t_2 within the domain, and for any $\theta \in (0,1)$,

$$f\big((1 - \theta)t_1 + \theta t_2\big) > (1 - \theta)f(t_1) + \theta f(t_2)$$

Now let $t_1 = z$, $t_2 = 1$, and $\theta = \frac{z-h}{z-1}$. Since $f(1) = 0$, the inequality above

becomes

$$f\left(\left(1-\frac{z-h}{z-1}\right)z+\frac{z-h}{z-1}\right)>\left(1-\frac{z-h}{z-1}\right)f(z)$$

which can be simplified to

$$f(h)>\left(1-\frac{z-h}{z-1}\right)f(z)$$

Since $h=gz+1-g$, we have

$$f(h)>\left(1-\frac{z-h}{z-1}\right)f(z)=\left(1-\frac{z-gz-1+g}{z-1}\right)f(z)=gf(z)$$

Therefore we have

$$\frac{f(h)}{f(z)}>g$$

and hence

$$g<\frac{h-\frac{1}{h^{\alpha-1}}}{z-\frac{1}{z^{\alpha-1}}}$$

and hence $w'(z)<0$. So when $\alpha>1$, $w(p)$ is a decreasing function of z, which is the ratio x/y.

When $0<\alpha<1$, we can see that $f''(t)>0$, and $f(t)$ is convex on the domain. And we will have

$$g>\frac{h-\frac{1}{h^{\alpha-1}}}{z-\frac{1}{z^{\alpha-1}}}$$

as well as $w'(z)>0$. So when $0<\alpha<1$, $w(p)$ is an increasing function of the ratio x/y. ∎

Appendix 3.4 Analytical analysis of the case of binary choice tasks

To support our statements in Section 3.3.2, we again examine the CPT equivalent of our theory in the case of binary choices. The table below is the same as Table 3.5.

	State 1 pq	State 2 $p(1-q)$	State 3 $(1-p)q$	State 4 $(1-p)(1-q)$
Lottery A	x_1	x_1	x_3	x_3
Lottery B	x_2	x_3	x_2	x_3

However, unlike in the previous case where a whole CPT-equivalent weighting function can be derived, we can now only obtain the CPT-equivalent probability weights (because one choice does not reveal the decision maker's value of the lotteries). The following proposition shows the properties of these probability weights, with a focus on the domain of positive payoffs.

Proposition 3 *Consider an agent's choice between lotteries A= (x_1, p; x_3, 1-p) and B= (x_2, q; x_3, 1-q) where $x_1 > x_2 > x_3 \geq 0$ and 0<p<q<1. The following two statements are equivalent:*

(i) The agent has a utility function $v(x)$ as specified in Assumption 2, a reference point as specified in Assumption 1, and his preference over the two lotteries are determined by equations (3.2), (3.4), (3.5);

(ii) The agent has the same utility function $v(x)$ but behaves according to Cumulative Prospect Theory with $r = 0$ (i.e. status quo) and with the following weights w(p) and w(q) assigned respectively to outcome x_1 of lottery A and outcome x_2 of lottery B:

$$w(p) = p \cdot T_p \quad \& \quad w(q) = q \cdot T_q$$

where

$$T_p = \frac{qL_{1,2} + (1-q)L_{1,3}}{pqL_{1,2} + p(1-q)L_{1,3} + q(1-p)L_{2,3}}$$

$$T_q = \frac{pL_{1,2} + (1-p)L_{2,3}}{pqL_{1,2} + p(1-q)L_{1,3} + q(1-p)L_{2,3}}$$

and $L_{1,2} = (v(x_1) - v(x_2))^\eta$, $L_{1,3} = (v(x_1) - v(x_3))^\eta$ and $L_{2,3} = (v(x_2) - v(x_3))^\eta$.

The key message of this proposition is that in this context, whether and how much a probability is overweighted or underweighted depends not only on the probabilities and outcomes of the lottery being valued, but also on the probabilities and outcomes of the alternative lottery in the choice set. For example, p is overweighted (underweighted) if $T_p > (<)$ 1, and the size of T_p depends on all the variables of the choice problem: x_1, x_2, x_3, p, and q. Small probabilities now need not be overweighted because a small p does not necessarily lead to $T_p > 1$. The similar analysis applies for q. We can roughly examine how these variables affect the probability weights with some assumption about the value of η. For simplicity, we assume $\eta = 1$ (which has been shown to give an inverse-S shaped probability weighting in valuation tasks with typical values of α and λ). Then it can be shown that:

1. T_p is a decreasing function of q and an increasing function of $L_{1,2}$. This means that, ceteris paribus, when x_2 decreases (getting closer to x_3) or q decreases (getting closer to p), probability p is more and more likely to be overweighted.

2. T_q is a decreasing function of p and an increasing (decreasing) function of $L_{1,2}$ if $p > (<)$ 0.5. For similar reasons, the larger p is, the less likely that q gets overweighted. In addition, given x_1 and x_3, as $L_{1,2}$ increases, $L_{2,3}$ has to decrease. Since probability q is related to the two states with outcome pairs (x_1, x_2) and (x_2, x_3), its weight will be increasing in $L_{1,2}$ only if the increase of $L_{1,2}$ has greater effect on T_q than the decrease of $L_{2,3}$. This requires $p > 0.5$.

From Proposition 3 we can also roughly identify conditions for T_p and T_q that are consistent with different shapes of the underlying probability weighting function. For example, consider only four types of the weighting function: Inverse-S type, S type, Overweighting type (i.e. all probabilities are overweighted) and Underweighting type (all probabilities are underweighted). Since $p < q$, if p is overweighted and q underweighted, we can think that the underlying probability weighting function is consistent with the Inverse-S type. Similarly, if p is underweighted and q overweighted, the underlying probability weighting function is consistent with the S type. If both are overweighted or underweighted, the underlying

probability weighting function cannot be identified. Specifically, we can derive the approximate conditions from this proposition for different types of probability weighting. This is shown in the table below, with the first row and first column listing the corresponding conditions.

Table 3.6: Approximate conditions for different types of implied probability weighting

	$L_{1,3} > 2qL_{2,3}$	$L_{1,3} < 2qL_{2,3}$
$p > 0.5$	Inverse-S type	not Overweighting type
$p < 0.5$	not Underweighting type	S type

In a typical choice between a small probability of winning large prize and large probabilities of small prize, the typical preference for the latter can be consistent with the S type probability weighting as p is small and q is large. For example, consider the preference-reversal-type choice between the lottery (£10, 0.4; £0, 0.6) and (£5, 0.8; £0, 0.2). The underlying weighting function for this choice will be consistent with the S type if $(v(10) - v(0))^\eta < 1.6(v(5) - v(0))^\eta$, which can hold with a concave utility function.

Appendix 3.5 Proof of Proposition 3

The table below is the same as Table 3.5.

	State 1 pq	State 2 $p(1-q)$	State 3 $(1-p)q$	State 4 $(1-p)(1-q)$
Lottery A	x_1	x_1	x_3	x_3
Lottery B	x_2	x_3	x_2	x_3

We begin with our theory, according to which, A is preferred to B if

$$\omega_1\big(v(x_1) - v(x_2)\big) + \omega_2\big(v(x_1) - v(x_3)\big) + \omega_3\big(v(x_3) - v(x_2)\big) > 0$$

which can be rearranged to be

$$(\omega_1 + \omega_2)v(x_1) + \omega_3 v(x_3) > (\omega_1 + \omega_3)v(x_2) + \omega_2 v(x_3) \tag{3.21}$$

Now consider the preference condition from Cumulative Prospect Theory. According to CPT, A is preferred to B if there is a probability weighting function $w(.)$ such that

$$w(p)v(x_1) + \big(1 - w(p)\big)v(x_3) > w(q)v(x_2) + \big(1 - w(q)\big)v(x_3) \tag{3.22}$$

For condition (3.21) to be equivalent to (3.22), we have

$$\frac{w(p)}{1 - w(p)} = \frac{\omega_1 + \omega_2}{\omega_3} \quad \& \quad \frac{w(q)}{1 - w(q)} = \frac{\omega_1 + \omega_3}{\omega_2}$$

Solving for $w(p)$ and $w(q)$, we obtain

$$w(p) = \frac{\omega_1 + \omega_2}{\omega_1 + \omega_2 + \omega_3} = \frac{1 + \dfrac{\omega_2}{\omega_1}}{1 + \dfrac{\omega_2}{\omega_1} + \dfrac{\omega_3}{\omega_1}}$$

$$w(q) = \frac{\omega_1 + \omega_2}{\omega_1 + \omega_2 + \omega_3} = \frac{1 + \dfrac{\omega_3}{\omega_1}}{1 + \dfrac{\omega_2}{\omega_1} + \dfrac{\omega_3}{\omega_1}}$$

According to equation (3.3), we have

$$w(p)$$

$$= \frac{pq\big(v(x_1) - v(x_2)\big)^\eta + p(1 - q)\big(v(x_1) - v(x_3)\big)^\eta}{pq\big(v(x_1) - v(x_2)\big)^\eta + p(1 - q)\big(v(x_1) - v(x_3)\big)^\eta + (1 - p)q\big(v(x_2) - v(x_3)\big)^\eta}$$

$$w(q)$$

$$= \frac{qp\big(v(x_1) - v(x_2)\big)^\eta + q(1 - p)\big(v(x_2) - v(x_3)\big)^\eta}{pq\big(v(x_1) - v(x_2)\big)^\eta + p(1 - q)\big(v(x_1) - v(x_3)\big)^\eta + (1 - p)q\big(v(x_2) - v(x_3)\big)^\eta}$$

Now define $L_{1,2}$, $L_{1,3}$ and $L_{2,3}$ such that $L_{1,2} \equiv (v(x_1) - v(x_2))^\eta$, $L_{1,3} \equiv (v(x_1) - v(x_3))^\eta$ and $L_{2,3} \equiv (v(x_2) - v(x_3))^\eta$, we have

$$w(p) = \frac{pqL_{1,2} + p(1-q)L_{1,3}}{pqL_{1,2} + p(1-q)L_{1,3} + q(1-p)L_{2,3}}$$

$$w(q) = \frac{qpL_{1,2} + q(1-p)L_{2,3}}{pqL_{1,2} + p(1-q)L_{1,3} + q(1-p)L_{2,3}}$$

Rearrange the equations to be

$$w(p) = p \cdot \frac{qL_{1,2} + (1-q)L_{1,3}}{pqL_{1,2} + p(1-q)L_{1,3} + q(1-p)L_{2,3}}$$

$$w(q) = q \cdot \frac{pL_{1,2} + (1-p)L_{2,3}}{pqL_{1,2} + p(1-q)L_{1,3} + q(1-p)L_{2,3}}$$

and define

$$T_p \equiv \frac{qL_{1,2} + (1-q)L_{1,3}}{pqL_{1,2} + p(1-q)L_{1,3} + q(1-p)L_{2,3}}$$

$$T_q \equiv \frac{pL_{1,2} + (1-p)L_{2,3}}{pqL_{1,2} + p(1-q)L_{1,3} + q(1-p)L_{2,3}}$$

Proposition 3 is proved. ∎

Chapter 4: An Experimental Study of Outcome-sensitive Decision Weighting

4.1 Introduction

An enormous body of experimental and field evidence (for surveys see e.g. Starmer 2000, Wakker 2010) has indicated descriptive limitations of the Expected Utility Theory (EU). The most popular alternatives to EUT belong to the family comprising Prospect Theory (PT henceforth, Kahneman & Tversky, 1979) and Cumulative Prospect Theory (CPT henceforth, Tversky and Kahneman 1992). A convenient feature of these models is that they retain the classical 'dual' structure of EUT, in the sense of having two components capturing, respectively, attitudes towards outcomes and towards probabilities. Unlike EUT, they model subjectivity toward probability through a probability weighting function which can be non-linear. In CPT, they also distinguish between decision weights and probability weights, with the former depending on the latter as well as on ranks of outcomes. Probability weighting has received many interpretations, but since PT, it has been widely understood as reflecting the intrinsic human attitude towards, or psychophysics of, chance (Kahneman & Tversky, 1984; Gonzalez & Wu, 1999).

The probability weighting function has been extensively investigated, both theoretically (e.g. Diecidue et al. 2009; Prelec 1998) and empirically. The most typical and widely found empirical result is an inverse-S shaped probability weighting function at the aggregate level (Abdellaoui 2000; Lattimore et al. 1992; Tversky & Kahneman 1992). However, evidence has been accumulating that the probability weighting functions may also sometimes be non-inverse-S (e.g. Alarié

and Dionne 2001; Humphrey and Verschoor 2004; Harbaugh et al. 2002) and to be dependent on non-probabilistic factors, such as people's demographic characteristics (such as gender, e.g. Fehr-Duda et al. 2006; or age, e.g. Harbaugh et al. 2002), on their emotional state (e.g. Fehr et al. 2007), on the nature of outcome (whether it is gain or loss) (e.g. Abdellaoui 2000), on stake sizes (e.g. Fehr-Duda et al. 2010), and on the emotional or affective content of the payoffs (e.g. Rottenstreich and Hsee 2001). Apart from the distinction between gains and losses, none of these factors is suggested by CPT, and the effects of stake sizes are inconsistent with it, unless they conform to the specific rank-dependent formulation of CPT. These factors are only consistent with the psycho-physical mechanisms to the extent they are mediated through perceptual mechanisms.

This evidence challenges our common understanding of the decision weighting function as the intrinsic part of individual decision-making under uncertainty and can hardly reconcile with PT or CPT, if decision weights do not only depend on probabilities or ranks of outcomes. Indeed, Tversky and Kahneman (1992) have implicitly expressed their concern about this:

"*...despite its greater generality, the cumulative functional is unlikely to be accurate in detail. We suspect that decision weights may be sensitive to the formulation of the prospects, as well as to the number, the spacing and the level of outcomes. In particular, there is some evidence to suggest that the curvature of the weighting function is more pronounced when the outcomes are widely spaced.*"

Note that, in this passage, they use 'decision weights' instead of 'probability weights', because if the dual structure of the decision model is retained, the former is a more accurate description of the weight attached to the value of outcomes. They mention four non-probabilistic factors that they thought might systematically affect decision weighting, two of which will be our focus: outcome spacing and outcome level. Outcome spacing is the size of gaps, in the case of monetary outcomes, between possible outcomes of a gamble, and outcome level refers to the general stake level or the expected payoff level of a gamble.

The effect of these two factors on decision weighting has been indirectly shown

in some experimental studies[1]. However, we are aware of only a few studies that explicitly investigate the question of whether and how outcome level and outcome spacing affect decision weighting. For example, Etchart-Vincent (2004, 2009) studied this question for the domain of losses. She found that, compared to small-loss gambles, large-loss gambles appear to enhance probabilistic optimism (or lower the decision weight of the worst outcome), while larger outcome spacing tends to increase pessimism (or decrease the decision weight of the best outcome). Fehr-Duda et al. (2010) studied the effect of stake level and find that, in the domain of gains, decision weights of the best lottery outcome are generally smaller for high-stake gambles than for low-stake ones, indicating that the effect of stake levels on decision weighting may contribute to the well-known evidence that risk aversion increases with stake levels (e.g. Binswanger 1981, Holt & Laury 2002), They also find that in the domain of losses, stake level has little effect on decision weighting. While their studies suggest outcome spacing and levels may affect decision weighting, they didn't isolate the two effects from one another. The two factors interact with each other and both can affect decision weighting. In other words, the change of a lottery's outcome spacing (level) could be accompanied with its change in outcome level (spacing). To examine the effect of either of the two factors, we need to control for the other one.

The primary goal of this chapter is to investigate experimentally whether and how outcome level and spacing, in the domain of gain, affect decision-weighting. More importantly, we will separate the effect of outcome level on decision weighting from the effect of outcome spacing, with the help of a clean experimental control. If decision weights are found to depend on outcome level and outcome spacing, it lends support to the theory presented in Chapter 3, in which Proposition 1 predicts the implied probability weighting function to be outcome-sensitive. However, please note that this experiment is *not* primarily designed to test our theory or to distinguish it from other theories of risk, although our experimental data would allow for a convenient test of one of the implications of Proposition 1 from the theory of Chapter

---

[1] See for example: Allais 1988, p. 243; Camerer 1992, p. 237; Harless and Camerer 1994, p. 1282.

3, which predicts that the effect of outcome spacing on decision weights vary with the utility curvature of the decision maker (see Appendix 3.3). In Section 4.4.4 we present the result of this test.

For simplicity, we consider only two-outcome lotteries in the form of {*x*, *p*; *y*, 1-*p*} with *p* being a probability, *x* and *y* being monetary outcomes, and *x*>*y*>0. Particularly we investigate how the decision weight attached to the best outcome *x* vary with lottery outcome level and outcome spacing.

In this chapter, outcome level of a lottery is defined as its expected payoff level, and outcome spacing of a lottery is defined as the *relative* outcome difference, i.e. the ratios of its best outcome to its worst outcome *x/y*. We investigate relative rather than absolute outcome spacing because the theory presented in Chapter 3 has specific predictions regarding how relative outcome spacing affects decision weighting. In addition, it seems more intuitive that relative rather than absolute outcome spacing is more likely to affect decision weighting. For example, consider two lotteries {£11, 0.3; £1, 0.7} and {£1000, 0.3; £990, 0.7}, which have the same absolute outcome spacing but dramatically different relative outcome spacing. A gap of £10 would probably be perceived as much more salient in the first lottery than the latter. We think this is likely to result in that people attach different weights to the best outcome of the first lottery and to that of the second lottery, even if outcome level has no effect on decision weighting. In contrast, PT and rank-dependent utility theories including CPT would predict that for lotteries {*x*, *p*; *y*, 1-*p*} with *x*>*y*, the decision weights attached to *x* should be independent of both *x* and the ratio *x/y*, provided *p* is held constant and variation in *x* and *y* does not affect either their ranking or their signs.

We report an econometric analysis of the data from a new experiment designed to test this prediction and a contrary prediction of the theory presented in Chapter 3. We apply econometric methods to estimate the behavioral model parameters of subjects, with a focus on the estimated decision weights attached to outcome *x*. Our main experimental finding is that, in general, outcome spacing has a systematic effect on decision weighting at both aggregate and individual levels, whereas outcome level

only has a systematic effect at individual level. For a given probability p, the estimated decision weights on x are generally smaller for lotteries with larger outcome spacing, and also smaller for lotteries with higher outcome levels. The theory from Chapter 3 is consistent with our data, but no strong conclusion can be drawn from our results about the implication of relations between utility curvature and the effect of outcome spacing.

The remainder of the chapter is organized as follows. Section 4.2 describes the experiment. Section 4.3 explains our theoretical framework and data analysis strategy. Results are reported and discussed in Section 4.4. The last section concludes.

4.2 Experimental Design

A key step in our experimental design is construction of four sets of lotteries, two of which differ in outcome levels and the other two in outcome spacing. We elicit each subject's Certainty Equivalents (CEs) for all lotteries in these sets. The four sets of lotteries have the following features[2]:

Set 1, with high average expected payoffs and high average x/y ratios;

Set 2, with low average expected payoffs and high average x/y ratios;

Set 3, with high average expected payoffs and low average x/y ratios;

Set 4, with low average expected payoffs and low average x/y ratios.

Table 4.1: The basic design structure

	High Payoff	**Low Payoff**
High Spacing	Set 1 (8 lotteries)	Set 2 (8 lotteries)
Low Spacing	Set 3 (8 lotteries)	Set 4 (8 lotteries)

---

[2] For a full list of the lotteries see Appendix 4.1.

As shown in Table 4.1, Set 1 and 2 together make Set HS (High Spacing), Set 3 and 4 together make Set LS (Low Spacing), Set 1 and 3 together make Set HP (High Payoff), and Set 2 and 4 together make Set LP (Low Payoff). To examine the effect of outcome spacing, we estimate subjects' preference parameters with HS lotteries and LS lotteries respectively, and compare the results. Similarly, by comparing preference parameters estimated with HP and LP lotteries we can know the effect of outcome level on decision weighting.

While the HS (HP) lotteries have generally higher x/y ratios (expected payoffs) than LS (LP) lotteries, it is important that there is also variability, with respect to both the payoff ratio and payoff level, for lotteries within each set, mainly because of the data analysis strategy we adopt. The reason for this will be better explained after we have introduced our econometric methods in the next section. As mentioned before, a key feature of our design is the capability to isolate the potential effect of outcome spacing from that of outcome level. To achieve this, we define Sets HP and LP in such a way that they have *statistically indistinguishable distributions* of lottery payoff ratio x/y, and define Sets HS and LS in such a way that they have *statistically indistinguishable distributions* of lottery expected payoffs. In this sense we can investigate the effect of either of the two factors while controlling for the other.

Table 4.2 shows that, statistically, HS and LS lotteries only differ in their distributions of the ratio x/y but not in their distributions of expected payoffs, and that statistically, HP and LP lotteries only differ in their distributions of expected payoffs but not in their distributions of ratio x/y. The technique we used is Mann-Whitney two-sample test with the null hypothesis that the two samples have the same distribution[3].

---

[3] For details of the statistics of the x/y ratio and expected payoffs of lotteries of each set, see Appendix 4.1.

Table 4.2: P-values of hypothesis tests regarding properties of different sets of lotteries

Lottery Probability	On ratio *x/y*		On expected payoff	
	H_0: HS=LS	H_0: HP=LP	H_0: HS=LS	H_0: HP=LP
p=0.09	0.000	0.993	0.993	0.000
p=0.21	0.000	1.000	1.000	0.000
p=0.33	0.000	0.993	0.993	0.000

Again, all lotteries are in the form of {*x, p*; *y*, 1-*p*} with *x>y>*0. We use three probabilities: *p*=0.09, 0.21, 0.33, and for each probability there are 32 lotteries divided into sets in the way shown in Table 4.1 as discussed above. We chose small probabilities spread across the range that previous studies have shown most likely to be overweighted, mainly because according to literature, the decision weights of these small probabilities seem to be more sensitive to lottery outcomes than moderate or large probabilities (e.g Etchart-Vincent 2004, Fehr-Duda et al 2010, Vieider et al. 2013). We had three probability levels primarily because we want to introduce some probability variance to the lotteries to prevent subjects potentially from forming some very simple decision heuristics, which may arise if they see only a single probability throughout the experiment. In addition, the variance in probability levels can also serve as a robustness check for our results. On the other hand, since our design is within-subject, we did not want to have too many different probabilities because each extra value of *p* requires another 32 CE tasks for subjects to complete and we don't want subjects to face too many tasks.

It is worth mentioning that we have spent considerable efforts on simulation tests to determine the details of our design and to ensure that our design works well with our econometric methods. Specifically, simulation tests helped us to arrive at suitable design choices, such as the number of lotteries in each set, the range of the *x/y* ratios and expected payoffs in each set, and the extent to which the HS (HP) and LS (LP) sets differ in their distributions of the *x/y* ratios (expected payoffs). To be more precise, these design choices are made such that our econometric methods can recover or back out the preference parameters of interest with simulated subjects' CE

data[4]. The current version of design we use has the highest chance to achieve this goal among all the versions we have tried[5].

The experimental task for subjects is to report their CEs of each of the 96 lotteries[6]. Figure 4.1 shows the type of task subjects saw. Lotteries were displayed on the left as a pie chart, and to make the task as simple as possible to subjects, we let them make their decisions by entering their CE of the lottery into the textbox on the right. They can enter any money amount (i.e. in pound and pence) between the lowest and highest money payoffs of the corresponding lottery. The order of the 96 tasks was randomized for each subject.
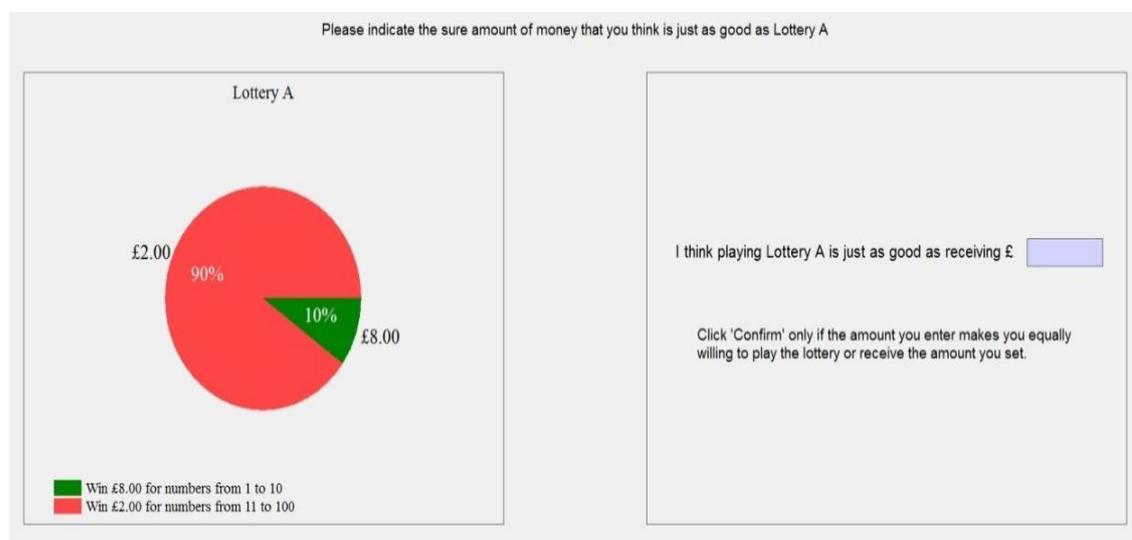
Figure 4.1: A sample CE-task screen

For each subject, the CE task for one randomly selected lottery was payoff-relevant. To incentivize subjects to truthfully report their CEs, we asked subjects to pick an envelope from a box of sealed envelopes before the experiment started, and told them that the envelope contains the information about their

---

[4] We define 'recover' as follows: the preference parameters are recovered if the distributions of the estimated parameters (over simulated virtual subjects) are statistically indistinguishable from the true parameter distributions, i.e. the parameter distributions we simulated. The simulations of virtual subjects were based on assumptions of normally distributed behavioral parameters, in our case, typical CPT model parameters.

[5] With each version of design, we simulated 40 samples each containing 40 subjects, applied our econometric method, and checked how many samples out of the 40 would pass our test. A sample passes our test if all parameter distributions we have simulated are recovered with this sample. The current version of the design had the best performance (26 out of 40 samples passed the test).

[6] As can be seen from Figure 4.1, we asked subjects to state the amount they see just as good as the lottery, instead of using the term Certainty Equivalent. Incentives are explained below.

payoff-relevant lottery and about an offer of an amount of sure money. Subjects are told that the payoff-relevant lottery could be any of the 96 lotteries they would see in the experiment, and the amount of offer could be any between the lowest and highest prizes of the payoff-relevant lottery. They were then told that they will play the payoff-relevant lottery if their reported CE of this lottery is larger than their offer of sure money, and they will receive the offer if their stated CE is equal to or less than their offer. The whole payoff determination process was carefully explained in the experimental instructions (see Appendix 4.4) to make sure subjects understood it clearly. In addition, we also introduced a quiz before the main tasks to test subjects' understanding of the tasks and the incentive scheme, and they were not allowed to start the main tasks unless they have answered all the quiz questions correctly.

The experiment was run at the CeDEx laboratory at the University of Nottingham. 96 student subjects participated in this experiment with an average payoff of £18.20 including a show-up fee of £3. They were placed into five separate sessions. For each session, subjects first picked an envelope from a full set of 96 (corresponding to the 96 lotteries in the experiment). Within each session, however, subjects picked the envelopes one by one without replacement, and at the same time they were told not to open the envelope until instructed. Then they read the instructions, completed the quiz, and proceeded to the experimental tasks. There was no time limit for completing a single task or all tasks. After all subjects have finished their tasks, they were asked to open their envelope and enter the indicated lottery code and amount of offer into the computer[7]. The program would retrieve subjects' decisions and show them their final payoff. For those who were to play the lottery, they were then offered to draw a chip from a bag of 100 chips, numbered from 1 to 100. The association between chip numbers and monetary outcomes was given to subjects in the initial description of the lottery, as the left hand panel of figure 4.1 illustrates. Each session took approximately an hour.

---

[7] The experimenter would check whether subjects have entered their code and offer truthfully before they received their payment. All subjects did enter the information truthfully.

4.3 Data Analysis Strategy

Our theoretical framework is more general than CPT, in the sense that decision weights are not restricted to be only probability-dependent or rank-dependent. Consider a lottery (*x*, *p*; *y*, 1-*p*) with *x* and *y* being outcomes (*x>y>*0) and *p* being a probability. Let the value of such lottery be represented by

$$V = w_x v(x) + (1 - w_x)v(y) \tag{4.1}$$

where w_x is the decision weight of the best outcome *x*, and $v(.)$ is an increasing utility function of monetary outcomes.

The econometric method we use is semi-parametric estimation, which has been widely used in preference elicitation studies such as Viscusi and Evans (2006) and Abdellaoui et al. (2011a). In other words, we estimate point decision weights along with utility parameters without assuming a parametric form of the weighting function. We choose this method mainly for two reasons. First, as mentioned earlier, we focus on specific small probabilities, so we don't require the whole probability-weighting function. Secondly, our core interest lies in comparing the estimated sizes of decision weights between HS and LS lotteries, and between HP and LP lotteries, and estimating decision weights directly would allow us to make comparisons of decision weights more straightforwardly, and probably more precisely than inferring them from estimated parameters of an assumed weighting function.

Our basic econometric model is represented by equation (4.1), with a power utility function $v(m) = m^\alpha$, where *m* is a monetary outcome. The power utility function is probably the most popular parametric family for fitting utility (Wakker, 2008). The advantage of a power utility function is that it can be conveniently interpreted and has also turned out to be the best compromise between parsimony and goodness of empirical fit in the context of prospect theory (Stott 2006). In addition, some studies have shown that for decision contexts of small and moderate

payoffs, the power utility function fits no worse than alternative classes of functions (e.g. Abdellaoui et al. 2008) such as the exponential or the 'expo-power' class first proposed by Saha (1993)).

So for binary lotteries, (x, p; y, 1-p) with $x>y>0$, we estimate the following model for each given p and for each of the four sets of lotteries (HS, LS, HP and LP):

$$ce = \widehat{ce} + \epsilon \tag{4.2}$$

where

$$\widehat{ce} = (wx^\alpha + (1-w)y^\alpha)^{1/\alpha} \tag{4.3}$$

For convenience we use w to denote the decision weight on the best lottery outcome x. The right hand side of equation (4.3) gives the CE obtained by applying the appropriate indifference condition to equation (4.1).

It is a standard practice to assume that individuals' decisions are noisy (e.g. Loomes 2005). There may be different sources of error, such as carelessness, hurry or inattentiveness, and in our setup, imprecise preferences (Hey and Orme, 1994). Moreover, as Harbaugh et al. (2010) suggest, failing to account for decision errors may significantly affect estimated risk preference parameters in ways that are task-dependent. To be specific, while decision errors are usually assumed to be normally distributed with zero mean, the error variance is more likely to be task-dependent. Given the nature of our elicitation task, we allow for two different sources of decision error heteroskedasticity[8]. First, each individual has to consider 96 lotteries with various payoff ranges. Since the observed certainty equivalent *ce* is stated by subjects and is forced to be between the largest and smallest payoffs of the lottery, it is very likely that the error is proportional to the payoff range. Secondly, since our approach models a representative agent's behavior, an individual's decision will most likely depart from the average prediction. As subjects may be heterogeneous with respect to their ability to find their true certainty equivalent, we allow, and expect, the error variance to differ by individual. Our error specification is

---

[8] Many studies have used heteroskedastic decision errors, such as Hey (1995), Blavatskyy (2007), and Fehr-Duda et al. (2010). In addition, Loomes (2005) has discussed reasons why it might be too restricted for models that have a stochastic component to impose constant decision error variance.

therefore:

$$\epsilon \sim N(0, (x-y)\sigma^2)$$

where the error variance, $(x-y)\sigma^2$, is proportional to the payoff range of lotteries.

To sum up, our data analysis strategy is to estimate equations (4.2) and (4.3), with respect to three parameters: α (power utility), w (decision weight of the best outcome) and σ (payoff-range-independent error parameter[9]). For each individual and each probability p=0.09, 0.21, 0.33, we have data comprising 32 CE values, and we did four estimations: one with HS lotteries, one with LS lotteries, one with HP lotteries and one with LP lotteries. By comparing the parameter estimates from the HS group and the LS group, we can see how relative outcome spacing (or the ratio x/y) of lotteries affect decision weights, and by comparing the parameter estimates from the HP group and the LP group, we can see how the expected payoff level of lotteries affect decision weights. In section 4.4, we report both aggregate-level analysis, in which we pooled all the individual data together and performed the same four estimations for each of the three probabilities, and individual-level analysis.

The estimation technique we use is Maximum Likelihood Estimation (MLE). Unlike Fehr-Duda et al. (2010) or Bruhin et al. (2010), we didn't use the finite mixture regression model to account for individual heterogeneity for two reasons. First, the mixture model is very complex, and as they discussed, various problems may be encountered when maximizing the likelihood function of a finite mixture regression model, and it is normal that not all of these problems can be solved. Secondly, they used the finite mixture regression model to account for non-CPT type behaviors, whereas our basic model is already more general than CPT, as explained earlier. Compatible with this more general theoretical framework is the use of semi-parametric method in which our estimated decision weights need not be consistent with CPT.

---

[9] For simplicity we will call σ the error parameter throughout the chapter.

4.4 Results

We first report aggregate level results, then individual level results. For the latter, we also report and discuss the results of a robustness check. Finally we discuss the results regarding the test of the implication of my theory from Chapter 3. As we consider only two outcome lotteries (*x*, *p*; *y*, 1-*p*) with *x>y,* we will simply use 'decision weight' to refer to 'the decision weight attached to outcome *x*' in this section.

4.4.1 Aggregate Level Results

Before we present our econometric results, we first report a general, and model-free, indicator of our subjects' revealed risk attitudes. The Relative Risk Premium index, RRP=(CE-EV)/EV, is calculated for each CE data point we obtained. Notice that slightly different from most of the literature (e.g. Dyer & Sarin 1982, Bruhin et al. 2010) we have defined RRP so that a positive value corresponds to risk-seeking behavior and a negative one to risk-averse behavior. Although this is unusual, it is convenient in the light of the results to be reported.

Generally subjects are risk-seeking, consistent with the fourfold pattern of risk attitudes predicted by Prospect Theory with an inverse-S shaped probability weighting function, that people are risk seeking for small probabilities of gains. They are more risk-seeking for HS lotteries than for LS lotteries, less risk-seeking for HP lotteries than for LP lotteries, and less risk seeking as probability of the best outcome increases.
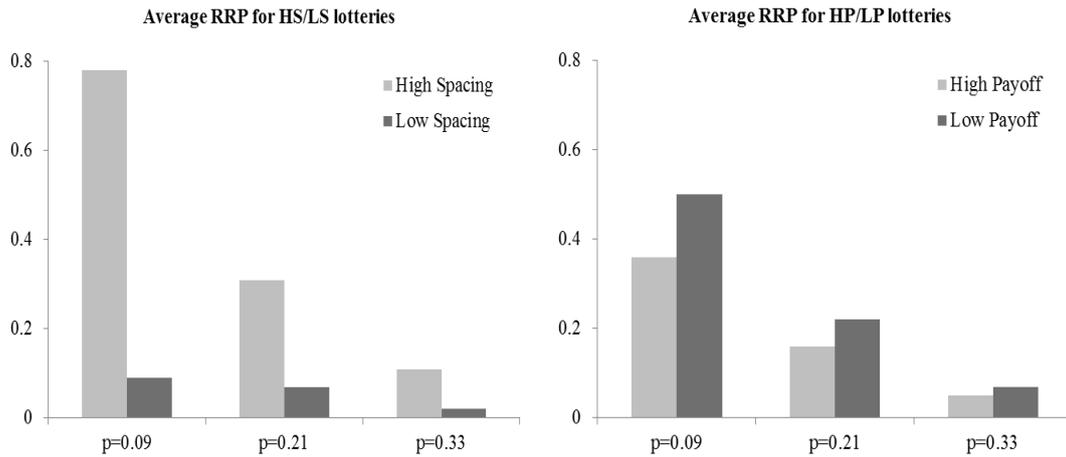
Figure 4.2: Average RRP for different probabilities and different sets of lotteries

Figure 4.2 shows the average RRP for all the 96 subjects over all lotteries of a given set and a given probability level. Since on average subjects are risk-seeking (according to this measure), it is unsurprising that people are more risk-seeking for HS lotteries than for LS lotteries, as shown on the left panel of Figure 4.2, because lotteries with high spacing have a larger *x/y* ratio and can be seen as more risky than lotteries with low spacing, when the probabilities of getting *x* are the same. However, the size of gap between RRP of HS lotteries and RRP of LS lotteries are surprisingly large and statistically significant (two-sided t-test, for all three probabilities p-values are 0.000). We will show later that this is probably because most subjects exhibited convex utilities of money in our study. The right panel of Figure 4.2 shows that people are less risk-seeking when the expected payoffs of the lotteries are larger. The difference is significant for p=0.09 and 0.21 (two-sided t-test, p-values are 0.001 and 0.024 respectively). This is consistent with the well-known finding (e.g. Kachelmeier & Shehata 1992, Beattie & Loomes 1997, Holt & Laury 2002) that people are less risk-seeking or more risk-averse when the stakes of bets are higher.

We now turn to one of our major interests, namely, whether the parameter estimates are different for different sets of lotteries.

Result 1: At aggregate level,

(i) The estimated curvatures of the utility functions are not significantly different

between HS/LS lotteries or between HP/LP lotteries (for the majority of the cases);

(ii) There is a tendency for each of the three probabilities to be overweighted. Decision weights are systematically[10] lower for HS lotteries than for LS lotteries, but are not systematically different between HP/LP lotteries.

(iii) Decision error parameter σ is systematically larger for HS and HP lotteries than for LS and LP lotteries.

Table 4.3 shows the parameter estimation results with all subjects' CE data pooled together. From this table we can see that, in most cases, the 'representative' subject exhibits convex utility functions, and in almost all cases, probabilities can be seen as overweighted, as all decision weights $w(p)$ are larger than their corresponding probability p. For each probability, we performed the Likelihood-Ratio test for each of the six pairs of parameters (for details see Appendix 4.2). According to the test results, utility curvature is not significantly different for HS and LS lotteries when p=0.09 or 0.21, and is not significantly different for HP and LP lotteries when p=0.21 or 0.33. Decision weights are significantly lower for HS than for LS lotteries when p=0.21 or 0.33, indicating a significant and systematic effect of outcome spacing on decision weighting.

Somewhat surprisingly, the effect of outcome level on decision weighting is neither significant nor systematic at aggregate level. One possibility is that our manipulation of expected payoffs is not strong enough. In other words, if we enlarge the difference of expected payoffs between HP and LP lotteries, we may observe significant effect of outcome level on decision weighting. However, our study also brings another possibility: since existing studies of the effect of outcome level (mainly Etchart-Vincent 2004 and Fehr-Duda et al. 2010) did not explicitly control for potential effect of outcome spacing, their finding of a significant effect of outcome level on decision weighting might be driven by an effect of outcome

---

[10] In our study, we say that an effect is systematic if for all three probability levels, the effect has the same direction, and for at least one of the three probabilities, the difference between parameter estimates is statistically significant. For example, the effect of outcome spacing is systematic because for all probability levels, w_{HS} is smaller than w_{LS}, and for p=0.21 and 0.33, w_{HS} are significantly different from w_{LS}. According to this definition, the effect of outcome level on decision weighting is not systematic at aggregate level.

spacing rather than outcome level.

Table 4.3: Pooled estimation of the parameters for different sets of lotteries

	Effect of Outcome Spacing				**Effect of Payoff Level**		
	p=0.09	*p*=0.21	*p*=0.33		*p*=0.09	*p*=0.21	*p*=0.33
α_{HS}	1.258	1.317	**1.369****	α_{HP}	**0.917****	1.021	1.147
	(0.164)	(0.175)	(0.183)		(0.100)	(0.097)	(0.103)
α_{LS}	1.399	1.459	**0.782****	α_{LP}	**1.385****	1.016	1.082
	(0.257)	(0.500)	(0.474)		(0.165)	(0.135)	(0.143)
w_{HS}	0.262	**0.307****	**0.352****	w_{HP}	0.300	0.359	0.393
	(0.030)	(0.033)	(0.037)		(0.021)	(0.020)	(0.020)
w_{LS}	0.291	**0.378****	**0.461****	w_{LP}	0.264	0.378	0.391
	(0.036)	(0.038)	(0.042)		(0.030)	(0.027)	(0.027)
σ_{HS}	**1.623****	**1.517****	**1.493****	σ_{HP}	**1.462****	**1.324****	**1.303****
	(0.029)	(0.027)	(0.027)		(0.026)	(0.024)	(0.023)
σ_{LS}	**0.479****	**0.499****	**0.548****	σ_{LP}	**0.846****	**0.886****	**0.914****
	(0.030)	(0.028)	(0.028)		(0.030)	(0.029)	(0.028)

Note:

1. Sample size *N*=96, so that for each estimation we have 1536 (96*16) observations;

2. Standard error in parentheses;

3. The parameter subscripts *HS, LS, HP*, and *LP* denote respectively the parameter estimations obtained with the HS lotteries, LS lotteries, HP lotteries and LP lotteries;

4. For each given probability, there are six pairs of parameter estimates, (α_{HS}, α_{LS}), (w_{HS}, w_{LS}), (σ_{HS}, σ_{LS}), (α_{HP}, α_{LP}), (w_{HP}, w_{LP}), and (σ_{HP}, σ_{LP}); where the two elements of one of these pairs are statistically significantly different from one another at 5%, according to the Likelihood-Ratio test, this is denoted by ** attached to both elements.

There is also a strong pattern of decision error variance between the different

lottery sets. The decision error parameter σ is systematically and significantly larger for HS lotteries and for HP lotteries. Given that HS and HP lotteries generally have larger payoff ranges $(x - y)$ than LS and LP lotteries respectively, the actual decision error variance of CE, $(x - y)\sigma^2$, is even larger for HS and HP lotteries. It might be that our findings on σ reveal how people genuinely make mistakes under risk, but we cannot rule out that this result is partly a product of our design, in the sense that our elicitation method magnifies the payoff range effect on decision error variances. Although we are confident that our subjects did understand the experimental tasks and the incentive mechanism, it is arguable that for the purpose of preference elicitation, the valuation-based approach has its drawbacks (see e.g. Harrison 1992, Dave et al. 2010). It would therefore be interesting to know whether our results here are robust to the elicitation method, by using for example the choice-based approach instead of the valuation-based approach. We leave this for future research.

4.4.2 Individual Level Results

In the previous section we have only considered the results for the average decision maker. The parameter estimates of the pooled model may be misleading if there is substantial heterogeneity in the population. Therefore, we extend the analysis to account for individual heterogeneity in parameter estimates by applying the estimation strategy to each individual.

We first calculated individuals' average RRP over the 96 lotteries. Figure 4.3 shows its distribution. 82.3% of the subjects are risk seeking (with positive average RRP).
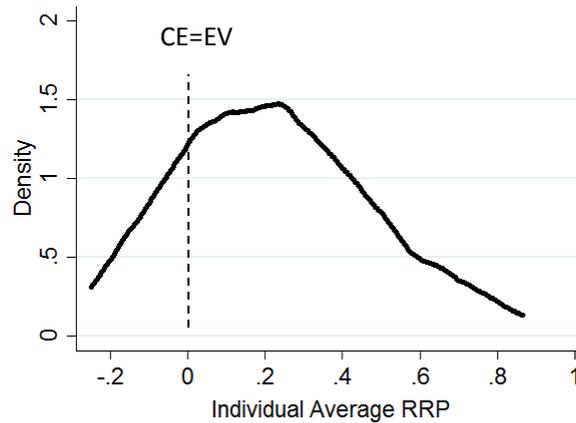
Figure 4.3: Distribution of individual average RRP over all subjects

To see the effect of outcome spacing and expected payoff level, we further calculated each individual's average RRP over lotteries of each of the four sets, HS, LS, HP and LP, pooling lotteries of different probabilities. The mean, median, and standard deviations of individual RRP for each set of lotteries are shown in Table 4.4.

Table 4.4: Statistics of individual average RRP for different lottery sets

	Mean	**Median**	**S.D.**
All Lotteries	0.29	0.25	0.32
HS	0.51	0.43	0.59
LS	0.06	0.05	0.06
HP	0.26	0.20	0.33
LP	0.32	0.27	0.32

Moreover, the distributions of individual average RRP are presented in Figure 4.4. Its left panel plots the distributions of individual average RRP for HS lotteries against LS lotteries, and the right panel shows the distributions of individual average RRP for HP lotteries against LP lotteries.
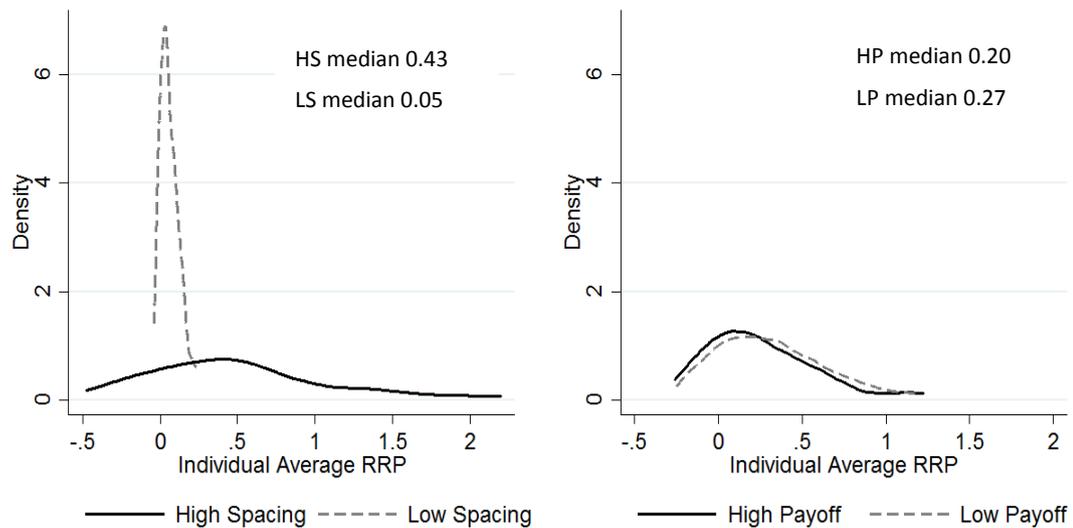
Figure 4.4: Distribution of individual average RRP for different sets of lotteries

Although by glancing at Figure 4.4 we can see that the effect of outcome spacing on RRP is much larger than the effect of expected payoff level, both pairs of distributions are significantly different at 5% according to the Wilcoxon Signed-Ranks test. The median of the average individual RRP is significantly larger for HS lotteries than for LS lotteries according to the sign test (p-value 0.000). We have checked whether this pattern also holds for risk averse subjects, and unsurprisingly the answer is no. For risk-averse subjects, their RRP is generally lower for HS lotteries than for LS lotteries. We can hence conclude that the overall result that people are more risk seeking for HS lotteries than for LS lotteries is mainly due to the fact that the majority of subjects in our study are risk seeking. The median of the average individual RRP is also significantly larger for LP lotteries than for HP lotteries (p-value 0.001). This is again consistent with the finding in the literature (e.g. Kuehberger et al. 1999, Holt & Laury 2002) that risk aversion (seeking) increases (decreases) with stake sizes.

In terms of individual heterogeneity, the variance of individual average RRP is remarkably larger for HS lotteries than for LS lotteries, whereas it is almost the same between HP and LP lotteries. This is not surprising because in our design, the *x/y* ratio is much larger for LS than for HS lotteries, and we forced subjects' reported CE

to be within the range (*y, x*). For example, for an LS lottery such as {£6, 0.21; £4, 0.79}, the largest possible RRP is approximately 0.36, whereas for an HS lottery such as {£50, 0.21; 10, 0.79}, the RRP can be as high as 2.72.

Now we report the parameter estimation results for each individual. Although 96 subjects participated in our experiment, 12 of them were excluded from the individual analysis, because no MLE results can be obtained for them[11]. Hence the following results are based on a sample size of *N*=84. For each individual and each probability *p*, we did four estimations, with respectively CE data for HS, LS, HP and LP lotteries. In general, the individual-level results regarding the parameters are qualitatively consistent with aggregate-level results, but the former give us additional information about distributions of parameters and about level of individual heterogeneity. These results can be summarized below.

Result 2: At individual level,

(i) There is considerable heterogeneity in individual parameter estimates.

(ii) Larger outcome spacing leads to smaller individual heterogeneity of utility curvature, a lower median decision weight, and both a larger median and greater individual heterogeneity of the decision error parameter.

(iii) Outcome level generally has no systematic effect on utility curvature or decision weighting, but higher expected payoffs lead to a systematically larger median and greater individual heterogeneity of the decision error parameter.

We will provide more details about the results of the parameter estimates one by one. Recall that for each individual we actually did 12 estimations, since we have three probabilities and for each probability we did four estimations with each of the four sets of lotteries (HS, LS, HP, LP). In what follows we will use $\tilde{\beta}_{\text{HS}}$, $\tilde{\beta}_{\text{LS}}$, $\tilde{\beta}_{\text{HP}}$, and $\tilde{\beta}_{\text{LP}}$ to denote individual-level parameter estimates respectively for the four different sets of lotteries, where $\beta \in \{\alpha, w, \sigma\}$ is one of the parameters of interest.

---

[11] We used STATA for the MLE, and no convergence could be achieved for those 12 subjects thus no estimation results were recorded by STATA.

Utility Parameter

The majority of subjects have convex utility functions for most of the estimations, as can be seen from Figure 4.5. It seems that this partly contributes to the risk-seeking behaviors of our subjects. There is much greater individual heterogeneity of the distribution of $\tilde{\alpha}_{LS}$ than the distribution of $\tilde{\alpha}_{HS}$, although the medians of $\tilde{\alpha}_{HS}$ and $\tilde{\alpha}_{LS}$ are only significantly different when $p=0.33$, according to Wilcoxon sign-test (p-value 0.000)[12]. Figure 4.5b shows the effect of expected payoff level on utility estimates. When $p=0.09$, the median $\tilde{\alpha}_{HP}$ is significantly smaller than the median $\tilde{\alpha}_{LP}$ according to the sign-test (p-value 0.000), but the differences are not significant when $p=0.21$ (p-value 0.380) or $p=0.33$ (p-value 0.445). Similar results hold for both risk-seeking and risk-averse subjects.
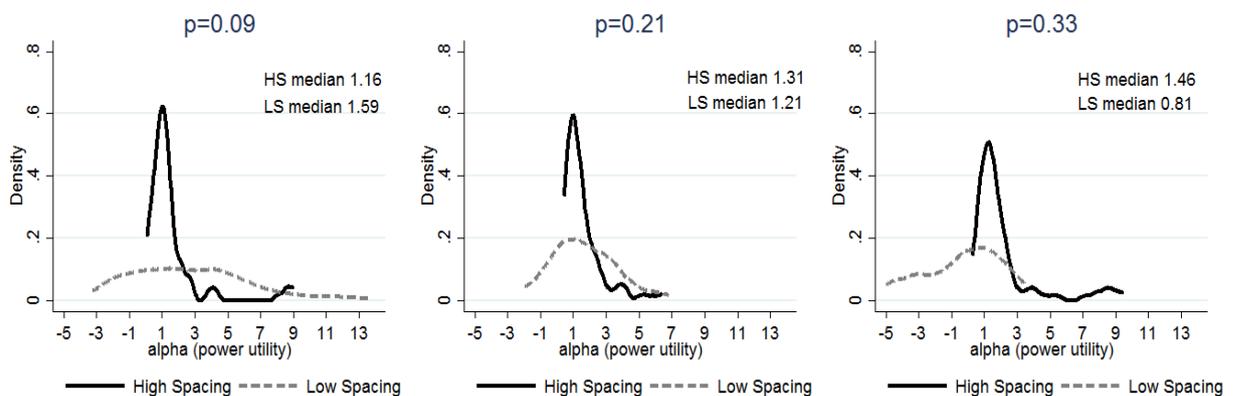
Figure 4.5a: Distributions of individual $\tilde{\alpha}_{HS}$ and $\tilde{\alpha}_{LS}$ for each probability

---

[12] For more details of the statistical tests of this section concerning medians and distributions of individual-level parameter estimates, see Table 4.14 of Appendix 4.3.3.
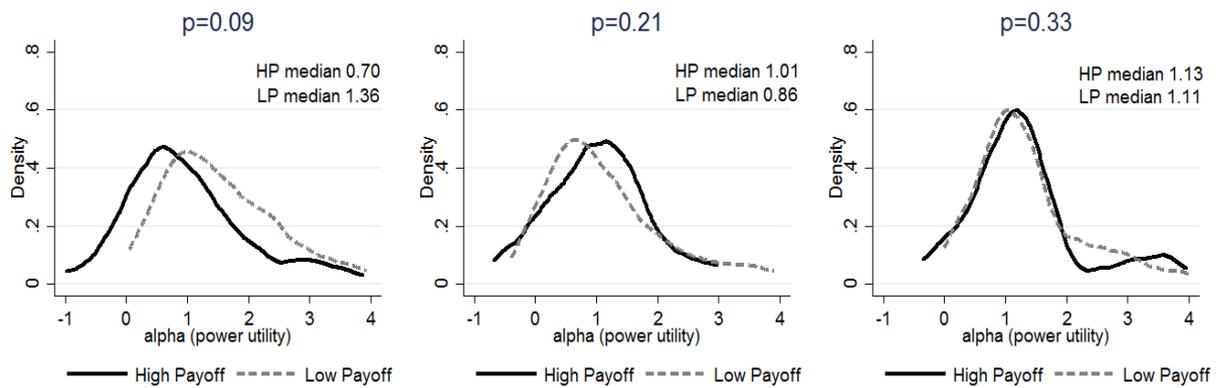
Figure 4.5b: Distributions of individual $\tilde{\alpha}_{HP}$ and $\tilde{\alpha}_{LP}$ for each probability

It is puzzling why there is much greater heterogeneity in the distribution of the estimated utility parameter for the LS lotteries than for HS lotteries, and with LS lotteries there are non-negligible negative utility parameter estimates. Such difference suggests that some of our subjects have substantially different utility estimates for HS and LS lotteries. There might be two explanations for this puzzling result. First, inconsistent with their decision attitude for HS lotteries, many subjects behaved more casually with LS lotteries, most probably because they cared less and paid less attention to their CE tasks when considering LS lotteries for which the best outcome and worst outcome do not differ much. For example, some of the subjects occasionally violated first-order stochastic dominance, that is, reported a higher CE for a (weakly) stochastically dominated lottery[13]. Since we have built in an error term to our model which is assumed to be normally distributed with zero mean, such occasional 'crazy' behaviors can hardly be captured by the error variance parameter if most of the errors are in the same direction. This may result in distortions of the estimated utility curvature, because for a given estimation, only x and y vary, but not the probability p, and lead to some psychologically implausible utility parameter estimates. Our data also shows that subjects who have negative $\tilde{\alpha}$ estimates violated dominance more frequently than those with positive $\tilde{\alpha}$ estimates.

The second explanation could be that subjects behaved sensibly and consistently

---

[13] These violations concern comparisons across CE tasks, not instances of CEs outside the range of payoffs of a given lottery.

with both HS and LS lotteries, but our estimation technique failed to recover their true parameters for LS lotteries. If the latter is true, it means that our estimation strategy probably cannot yield reliable, that is, unique and significant estimation results with binary lotteries that have a low *x/y* ratio. We did two things to try to rule out this possibility. First, as mentioned before, we did a simulation test to ensure that our estimation strategy works with our current lottery design, and the test shows that it works well. Secondly, we ran each estimations 20 times, (for both aggregate and individual level analysis), and can confirm that our estimation results are significant, and stable over the 20 replications.

Given that the first explanation is favored, there is a possible concern that the inconsistency of utility estimates across lotteries sets would bias the estimates of decision weights and the error parameter. In other words, the systematic and significant effect of outcome spacing on decision weights and on decision error variance may be results of the individual difference of utility curvatures between HS and LS lotteries. However, in a later section we show that this is not the case according to our robustness checks.

Decision Weight

In terms of decision weights (attached to the best outcome of a lottery), the majority of subjects overweight *p*=0.09 and 0.21, and about a half of subjects overweight *p*=0.33, as Figure 4.6 shows. For all three probabilities 0.09, 0.21, and 0.33, the median $\widetilde{w}_{\text{HS}}$ are significantly lower than the median $\widetilde{w}_{\text{LS}}$ according to the sign-test (p-values 0.001, 0.000, 0.000, respectively), whereas there is no systematic difference between the median $\widetilde{w}_{\text{HP}}$ and $\widetilde{w}_{\text{LP}}$. Specifically, only for *p*=0.09, are the median $\widetilde{w}_{\text{HP}}$ and $\widetilde{w}_{\text{LP}}$ differ significantly according to the sign-test (p-value 0.004), and the directions of the effect of expected payoff level are opposite for *p*=0.09 and *p*=0.21.
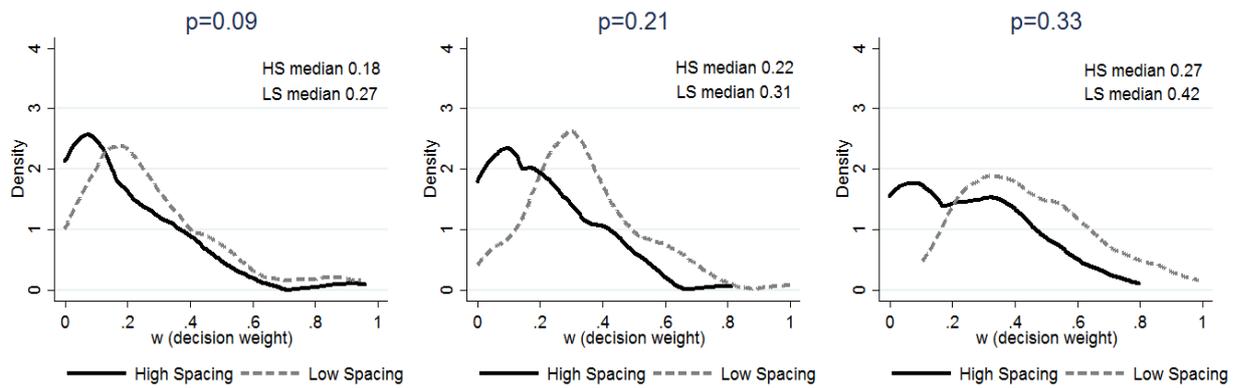
Figure 4.6a: Distributions of individual $\widetilde{w}_{HS}$ and $\widetilde{w}_{LS}$ for each probability
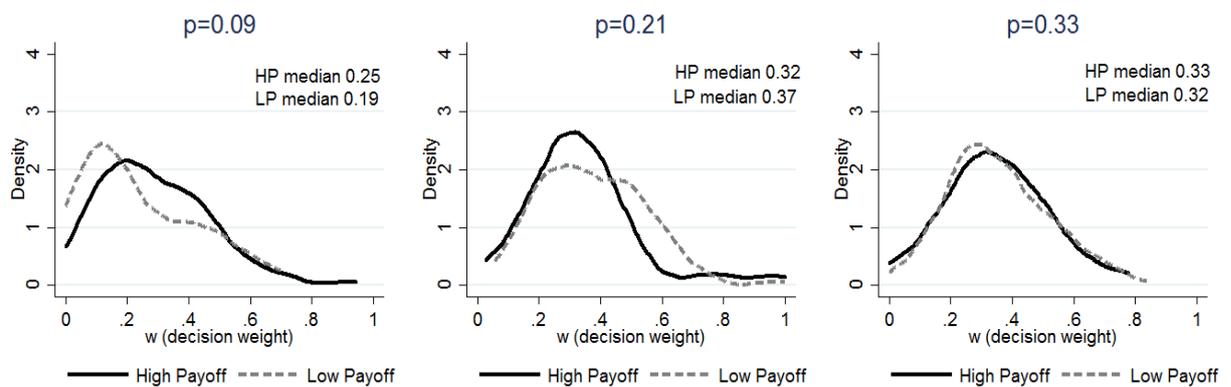
Figure 4.6b: Distributions of individual $\widetilde{w}_{HP}$ and $\widetilde{w}_{LP}$ for each probability

There is also no obvious difference in the level of individual heterogeneity across different sets of lotteries. Qualitatively similar results hold for both risk-seeking and risk-averse subjects.

It seems that decision weights have stronger patterns at individual-levels than at aggregate-level, and that outcome spacing has greater and more predictable impact on decision weighting than expected payoff level does. As we have discussed, this is unsurprising given that our manipulation of payoff level difference is not strong compared to the literature (e.g. Kachelmeier & Shehata 1992, Fehr-Duda et al. 2010).

What is more interesting is the direction of effect of outcome spacing. One might expect that the estimated w being higher for HS lotteries than for LS lotteries, because for the former the payoff gap $(x - y)$ is much larger, so that outcome x would become much more attractive in HS lotteries than in LS lotteries, and that

people would tend to overweight the weight attached to *x* more. However, our results show the opposite, that people exhibit significantly higher decision weights for LS lotteries. We will show in a Section 4.4.4 that this finding on decision weights can potentially be accommodated by the theory presented in Chapter 3, in which Proposition 1 implies that the direction of the effect of outcome spacing on decision weights can depend on the utility parameter α.

Decision Error Parameter

The effects on decision error parameter σ are stronger and more stable than on the other two parameters. Consistent with the aggregate-level results, at individual level, subjects have significantly larger median $\tilde{\sigma}_{HS}$ than $\tilde{\sigma}_{LS}$ (p-value 0.000 for all three probabilities), and significantly larger median $\tilde{\sigma}_{HP}$ than $\tilde{\sigma}_{LP}$ (p-value 0.000, 0.027, 0.021 respectively for p=0.09, 0.21, 0.33), according to the sign-test. There is also greater individual heterogeneity of the distributions of $\tilde{\sigma}_{HS}$ and $\tilde{\sigma}_{HP}$, as shown in Figure 4.7. However this result does not hold for risk-averse subjects, indicating that the gap shown in Figure 4.7 is mainly driven by risk-seeking subjects who are the majority.
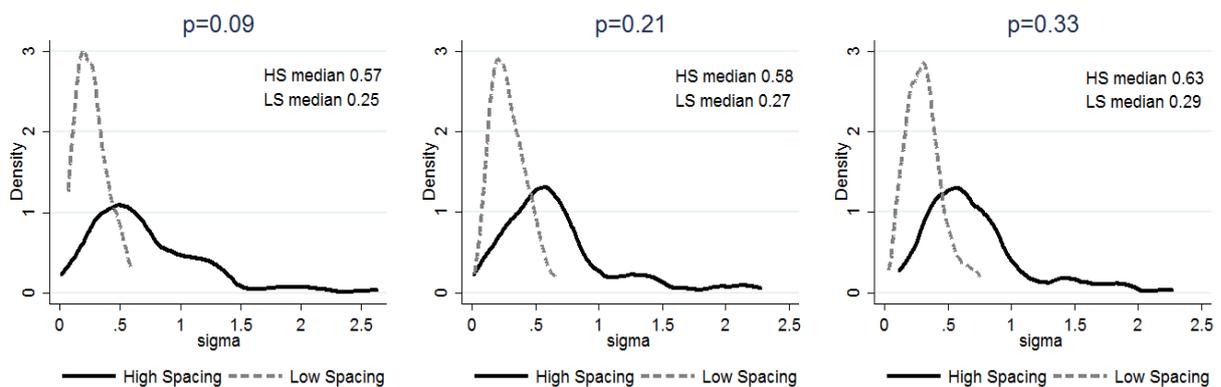

Figure 4.7a: Distributions of individual $\tilde{\sigma}_{HS}$ and $\tilde{\sigma}_{LS}$ for each probability
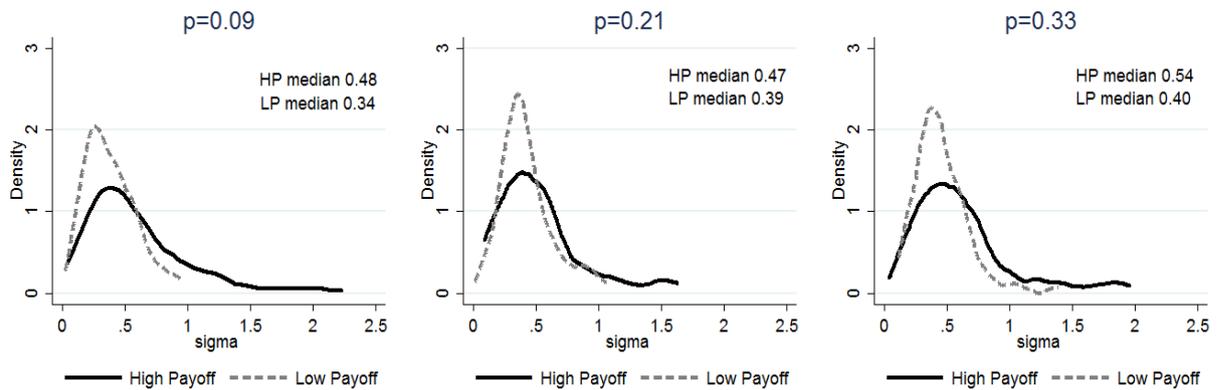
Figure 4.7b: Distributions of individual $\tilde{\sigma}_{\text{HP}}$ and $\tilde{\sigma}_{\text{LP}}$ for each probability

Again, since in our model specification, the error variance is correlated with lottery payoff ranges $(x - y)$, the difference of the actual decision error of CE would be much bigger, and the difference of the level of individual heterogeneity may be considerably larger for HS and HP lotteries than for LS and LP lotteries. In terms of heterogeneity, admittedly, this result may largely be due to the fact that subjects are forced to state their CEs between the range (y, x), and that HS lotteries have much wider payoff ranges than LS lotteries. However, this fact does not predict that people *will* make larger decision errors when they are allowed to. We conjecture that, as our results suggest, the possibility of larger mistakes itself induces a higher tendency to make larger mistakes. It is also interesting to ask whether this 'error boundary effect' is also significant beyond the type of elicitation task we use, but this is a question we defer to further research.

4.4.3 Robustness Checks

So far, we have found systematic impact of outcome spacing, but no systematic effect of outcome level, on decision weighting. We have also found that both outcome spacing and outcome level affect decision errors. However, as mentioned before, that some of our subjects have dramatically different utility parameter

estimates for HS and LS lotteries, we are concerned that this difference in utility estimates across lottery sets produces or magnifies the observed effects on decision weights and error variances. As a check on this, we investigate whether these effects would become smaller or even disappear if we hold each subject's utility parameter constant between HS/LS lotteries and between HP/LP lotteries. We highly doubt that the individual-level utility curvature estimates we obtained for LS lotteries, as shown in Figure 4.5a, reveal the subjects' true attitude towards money outcomes.

We did three robustness checks, with three ways to hold the utility estimates constant across HS/LS and across HP/LP sets.

Robustness Check 1: For each individual and each probability, we force $\tilde{\alpha}_{\mathrm{LS}}$ to be the same as $\tilde{\alpha}_{\mathrm{HS}}$, and $\tilde{\alpha}_{\mathrm{LP}}$ to be the same as $\tilde{\alpha}_{\mathrm{HP}}$. In other words, we re-estimate $\tilde{w}_{\mathrm{LS}}$ and $\tilde{\sigma}_{\mathrm{LS}}$ for the LS set with $\tilde{\alpha}_{\mathrm{HS}}$ taken as a constant utility parameter. Similarly, to test the outcome level effect, we re-estimate $\tilde{w}_{\mathrm{LP}}$ and $\tilde{\sigma}_{\mathrm{LP}}$ for LP lotteries with $\tilde{\alpha}_{\mathrm{HP}}$ taken as a constant utility parameter.

Robustness Check 2: Similar to the first check, except that we now force $\tilde{\alpha}_{\mathrm{HS}}$ to be the same as $\tilde{\alpha}_{\mathrm{LS}}$, and $\tilde{\alpha}_{\mathrm{HP}}$ to be the same as $\tilde{\alpha}_{\mathrm{LP}}$.

Robustness Check 3: For each individual and each probability, we estimate the three parameters with all the 32 lotteries, take the estimated utility parameter $\tilde{\alpha}_{\mathrm{ALL}}$ as constant across sets, and estimate w and σ respectively for HS, LS, HP, and LP lotteries.

With these checks, we implicitly assume that individual utility curvature is not affected by outcome spacing or outcome levels of lotteries. We are not aware of research on how outcome spacing or relative stake level affect utility curvature, but there has been a sizable body of evidence about the (absolute) stake level effect (e.g. Booij et al. 2010, Bombardini & Trebbi 2012, Vieider et al. 2013). These evidence, along with Rabin (2000)'s view, generally support the idea that people have changing relative risk aversion only when very large stakes are involved. Most of the studies that show evidence of an effect of payoff level on utility curvature used a payoff scaling-up factor of at least 10 (e.g. Scholten & Read 2014, Bouchouicha & Vieider 2016). Since in our design, the expected payoff level has only approximately doubled

for HP than for LP lotteries, we don't think our assumption about utility curvature would be too restrictive for the purpose of the robustness check. The results of these checks are summarized below (with details of these results provided in Appendix 4.3).

Result 3:

(i) All three robustness checks yield qualitatively similar results (to previous ones reported in Section 4.4.2) regarding the effect of outcome spacing on decision weights and decision error, and regarding the effect of outcome level on decision error.

(ii) Unlike the previous results, outcome level is now found to have a systematic effect on decision weighting. In these checks, decision weights are generally larger for LP lotteries than for HP lotteries.

The results confirm that the effect of outcome spacing on decision weighting as well as on the decision error parameter is strong and robust. Furthermore, the robustness checks update our finding about the effect of payoff level on decision weighting, which this is now consistent with the handful of literature (Etchart-Vincent 2004, 2009; Fehr-Duda et al. 2010; Vieider et al. 2013). This seems to suggest that, for the baseline estimation in which utility parameters are not held constant across sets of lotteries, the inconsistency of estimates of $\tilde{\alpha}_{\mathrm{HP}}$ and $\tilde{\alpha}_{\mathrm{LP}}$ (for some subjects) more or less distorts the results about the effect of outcome level on decision weights.

Table 4.5 shows how the effect of outcome level on decision weights differs across our estimation setups. As can be seen, in every column except those for $p = 0.09$, the four estimated decision weights seem remarkably stable. For all three robustness checks, the effect of expected payoff level becomes systematic. For all probabilities, $\tilde{w}_{\mathrm{HP}}$ is smaller than $\tilde{w}_{\mathrm{LP}}$, indicating that people seem to be more pessimistic about gambling when the stakes are higher. Moreover, according to

Wilcoxon sign-test, the gaps between median $\widetilde{w}_{HP}$ and $\widetilde{w}_{LP}$ are statistically significant at 5% for *p*=0.09 and 0.21 (p-values of these tests are all 0.00), though not for *p*=0.33.

Table 4.5: Medians of individual estimates of $\widetilde{w}_{HP}$ and $\widetilde{w}_{LP}$ for different setups

	p=0.09		**p=0.21**		**p=0.33**	
	$\widetilde{w}_{HP}$	$\widetilde{w}_{LP}$	$\widetilde{w}_{HP}$	$\widetilde{w}_{LP}$	$\widetilde{w}_{HP}$	$\widetilde{w}_{LP}$
Baseline Estimation	0.25	0.19	0.32	0.37	0.34	0.32
Robustness Check 1	0.25	0.31	0.32	0.36	0.34	0.37
Robustness Check 2	0.16	0.19	0.32	0.37	0.31	0.32
Robustness Check 3	0.23	0.29	0.32	0.36	0.32	0.36

Note: Again, baseline estimation refers to the one reported in Section 4.4.2 in which the utility parameter is not held constant across sets of lotteries.

On average, subjects seem to neither overweight nor underweight the probability of 0.33 in our study, and the effect of outcome level almost disappears in this case. This leads us to a conjecture that, as mentioned before, that the size of the effect of outcomes on decision weighting seems to vary across probability levels.

4.4.4 Testing the implication of Proposition 1 from Chapter 3

For simplicity we will call the theory presented in Chapter 3 TC3. In general, the results of decision weights we have shown so far stand against PT and CPT, but could be accommodated by TC3 which allows decision weighting to be sensitive not only to outcome ranks but also (relative) outcome spacing. Furthermore, TC3 also has a specific prediction, derived from Proposition 1 of Chapter 3 (Section 3.3.1), that for individuals with a concave utility function of monetary outcomes, their decision weights generally increase with outcome spacing (or, in our two-outcome

lottery case, with the *x/y* ratio), and for individuals with a convex utility function their decision weights would decrease with outcome spacing. It is important to notice that this prediction relies on three premises. First, preference parameters are measured using CPT with CE data. Secondly, the power utility function is assumed; Thirdly, individual power utility parameter is independent of lottery outcome spacing. Hence, the robustness checks make it convenient to test Proposition 1's implication about the effect of outcome spacing, with individual utility estimates being held constant across lottery groups.

Since we have three robustness checks, our test results differ slightly with respect to which robustness check is considered, that is, which utility parameter is taken as the outcome-spacing-independent utility parameter, $\tilde{\alpha}_{HS}$, $\tilde{\alpha}_{LS}$, or $\tilde{\alpha}_{ALL}$. For the first two robustness checks, we force the utility parameter to be outcome-spacing-independent, whereas for the third check, since the estimation set is the combination of HS and LS lotteries, $\tilde{\alpha}_{ALL}$ can be naturally seen as outcome-spacing-independent.

To test Proposition 1's implication, we first examine the proportions of subjects that behaved consistently with our prediction. For each probability *p*, we can sort subjects into four types according to their estimated utility function curvature and decision weight sizes[14]:

Type 1: $\tilde{\alpha} > 1$ and $\widetilde{w}_{HS} > \widetilde{w}_{LS}$

Type 2: $\tilde{\alpha} > 1$ and $\widetilde{w}_{HS} < \widetilde{w}_{LS}$

Type 3: $0 < \tilde{\alpha} < 1$ and $\widetilde{w}_{HS} > \widetilde{w}_{LS}$

Type 4: $0 < \tilde{\alpha} < 1$ and $\widetilde{w}_{HS} < \widetilde{w}_{LS}$

Among these, Type 2 and Type 3 are consistent with the prediction. So we calculated the proportion of subjects belonging to Type 2 or 3, with respectively the

---

[14] Since a decreasing utility function of money is neither realistic nor concerned with our theoretical prediction, we exclude those with negative power utility parameters from this test.

parameter estimates from the three robustness checks[15]. Our results show that the average proportions of subjects (over three probability levels) consistent with the prediction are respectively 74%, 85%, and 55% for Robustness Check 1, 2 and 3 respectively.

We also grouped subjects by whether they have concave or convex utility functions, and used the one-sided sign-test to test whether the median decision weight estimates differ significantly for HS and LS lotteries in the predicted direction. These results are summarized in Table 4.6. For subjects with concave utility, we tested the null hypothesis that $\widetilde{w}_{\text{HS}} = \widetilde{w}_{\text{LS}}$ against the hypothesis that $\widetilde{w}_{\text{HS}} > \widetilde{w}_{\text{LS}}$. For subjects with convex utility, we tested the same null hypothesis against the hypothesis that $\widetilde{w}_{\text{HS}} < \widetilde{w}_{\text{LS}}$. A p-value of less than 5% suggests that the alternative hypothesis is accepted, and hence our prediction is supported.

In Table 4.6, the hypothesis tests for which the p-values are denoted with ** lend support to Proposition 1's prediction. We can see that although the test results vary across the three checks, on average more than a half of the hypothesis tests favor our theoretical prediction. Particularly, these favoring tests correspond to the majority of our subjects who exhibit convex utility functions, while our prediction seems to work not very well with subjects who have concave utilities. Table 4.6c shows that for the latter type of subjects, the p-values are close to 1, indicating that the effect of outcome spacing on decision weights is insignificant in these cases.

However, although the results as shown in Table 4.6 do not completely fit the prediction derived from Proposition 1 of TC3, it is always possible that the prediction fails because the second or the third premise mentioned above doesn't hold. Furthermore, these results seem to suggest dependence of the effect of outcome spacing on the curvature of utility. Such dependence can play a role in TC3, but not in PT or CPT. Therefore, we think that TC3 is generally supported, given that we

---

[15] For example, with Robustness Check 1, Type 2 would be those with $\tilde{\alpha}_{\text{HS}} > 1$ and $\widetilde{w}_{\text{HS}} < \widetilde{w}_{\text{LS}}$, and Type 3 would be those with $0 < \tilde{\alpha}_{\text{HS}} < 1$ and $\widetilde{w}_{\text{HS}} > \widetilde{w}_{\text{LS}}$. Notice that the $\widetilde{w}_{\text{LS}}$ here are estimates from Robustness Check 1 but not from the baseline estimations. Since in Robustness Check 1 we have set $\tilde{\alpha}_{\text{LS}}$ as a constant equal to $\tilde{\alpha}_{\text{HS}}$, the estimated $\widetilde{w}_{\text{LS}}$ in Robustness Check 1 can be different from that in the baseline estimations. Similarly, in Table 4.6a, 4.6b, and 4.6c, the decision weight estimates are from Robustness Check 1, 2, and 3 respectively, but not from the baseline estimations.

Table 4.6a: Statistical tests for different types of subjects (Robustness Check 1)

Probability Group	Subject type (No. of subjects)	Alternative Hypothesis (p-value)
p=0.09	$\tilde{\alpha}_{HS} > 1$ (N=51)	$H_1: \tilde{w}_{HS} < \tilde{w}_{LS}$ (0.000**)
	$0 < \tilde{\alpha}_{HS} < 1$ (N=27)	$H_1: \tilde{w}_{HS} > \tilde{w}_{LS}$ (0.010**)
p=0.21	$\tilde{\alpha}_{HS} > 1$ (N=58)	$H_1: \tilde{w}_{HS} < \tilde{w}_{LS}$ (0.000**)
	$0 < \tilde{\alpha}_{HS} < 1$ (N=22)	$H_1: \tilde{w}_{HS} > \tilde{w}_{LS}$ (0.206)
p=0.33	$\tilde{\alpha}_{HS} > 1$ (N=60)	$H_1: \tilde{w}_{HS} < \tilde{w}_{LS}$ (0.000**)
	$0 < \tilde{\alpha}_{HS} < 1$ (N=17)	$H_1: \tilde{w}_{HS} > \tilde{w}_{LS}$ (0.314)

Table 4.6b: Statistical tests for different types of subjects (Robustness Check 2)

Probability Group	Subject type (No. of subjects)	Alternative Hypothesis (p-value)
p=0.09	$\tilde{\alpha}_{LS} > 1$ (N=49)	$H_1: \tilde{w}_{HS} < \tilde{w}_{LS}$ (0.000**)
	$0 < \tilde{\alpha}_{LS} < 1$ (N=9)	$H_1: \tilde{w}_{HS} > \tilde{w}_{LS}$ (0.000**)
p=0.21	$\tilde{\alpha}_{LS} > 1$ (N=44)	$H_1: \tilde{w}_{HS} < \tilde{w}_{LS}$ (0.000**)
	$0 < \tilde{\alpha}_{LS} < 1$ (N=19)	$H_1: \tilde{w}_{HS} > \tilde{w}_{LS}$ (0.179)
p=0.33	$\tilde{\alpha}_{LS} > 1$ (N=31)	$H_1: \tilde{w}_{HS} < \tilde{w}_{LS}$ (0.000**)
	$0 < \tilde{\alpha}_{LS} < 1$ (N=11)	$H_1: \tilde{w}_{HS} > \tilde{w}_{LS}$ (0.033**)

Table 4.6c: Statistical tests for different types of subjects (Robustness Check 3)

Probability Group	Subject type (No. of subjects)	Alternative Hypothesis (p-value)
p=0.09	$\tilde{\alpha}_{ALL} > 1$ (N=35)	$H_1: \tilde{w}_{HS} < \tilde{w}_{LS}$ (0.000**)
	$0 < \tilde{\alpha}_{ALL} < 1$ (N=37)	$H_1: \tilde{w}_{HS} > \tilde{w}_{LS}$ (0.997)
p=0.21	$\tilde{\alpha}_{ALL} > 1$ (N=39)	$H_1: \tilde{w}_{HS} < \tilde{w}_{LS}$ (0.000**)
	$0 < \tilde{\alpha}_{ALL} < 1$ (N=30)	$H_1: \tilde{w}_{HS} > \tilde{w}_{LS}$ (1.000)
p=0.33	$\tilde{\alpha}_{ALL} > 1$ (N=53)	$H_1: \tilde{w}_{HS} < \tilde{w}_{LS}$ (0.036**)
	$0 < \tilde{\alpha}_{ALL} < 1$ (N=21)	$H_1: \tilde{w}_{HS} > \tilde{w}_{LS}$ (0.987)

have found a significant effect of outcome spacing on decision weighting, but this specific prediction tested in this section is only partly supported. For Robustness Check 1 and 2, Proposition 1's prediction is supported in the majority of cases, but for Robustness Check 3, only a half of the cases are consistent with that prediction. Given the results, it seems still inconclusive whether the effect of outcome spacing on decision weighting depends on utility curvatures or not.

4.5 Conclusion and Discussions

This study pursues two goals. First, we investigate and isolate the effect of outcome spacing and outcome level on decision weighting. Our data suggest that, for small probabilities less than 1/3, outcome spacing has a significant effect on decision weighting at both aggregate and individual levels, whereas outcome level only has significant effect at individual level. For a given probability p, subjects generally exhibit smaller decision weights (of the best lottery outcome) for lotteries with larger outcome spacing, and smaller decision weights (of the best lottery outcome) for lotteries with higher expected payoff levels. The second goal is to test a prediction of the theory in Chapter 3 of this thesis. We find that this prediction is partly supported, but no firm conclusion can be drawn regarding whether decision weighting for risk is affected by utility curvatures.

Although our results only apply to small probabilities below 1/3, it is worth mentioning the potential theoretical and prescriptive implications of these results. First, from a theoretical point of view, the fact that decision weighting is systematically sensitive to outcome spacing and levels is a challenge to most of the theories of decision under risk that typically assume separability of decision weights and outcome valuation (e.g. PT). Decision models may misrepresent risk preferences considerably when decision weights interact with the level of payoffs or other lottery

characteristics in a material way. Although in CPT, decision weights can be different from probability weights[16], the dependence of decision weights on outcomes is perhaps restricted too much by that the only feature of the outcomes that matters for the weights is their ranking. In light of our evidence, we think that models that stay too close to the classical 'dual' structure (i.e. a component of decision weighting and a component of utility of outcomes) would be too simple to explain the evidence. But modifying the decision weighting function so as to incorporate more factors, such as outcome spacing and outcome level, would inevitably entail some loss of parsimony, and it would be even trickier to incorporate factors such as the length of delay of uncertainty resolution (Abdellaoui et al. 2011b), source of uncertainty (Abdellaoui et al. 2011a), or the type of lottery outcomes (Rottenstreich and Hsee 2001), all of which are found to have impact on decision weighting.

Secondly, since the majority of our subjects are risk-seeking in the experiment, our main finding that decision weights are generally smaller for lotteries with larger outcome spacing is counter-intuitive under the framework of CPT or other rank-dependent utility models (e.g. Quiggin 1982, 2012), even if decision weights can depend on outcome spacing. In other words, if the *w* parameter we have measured is understood as a decision weight *attached to the utility of consequences*, it is hard to understand why the decision weight of the best outcome of a lottery is lower when the best outcome is much better relative to its worst outcome.

We think our results have brought a serious question regarding how we should understand the concept of decision weight and whether the classical 'dual' structure of decision models is still helpful for us to understand and predict individual behaviors. After all, TC3 has provided a potential ground for accommodating these results. We speculate that other theories such as Loomes & Sugden (1982), Bell (1985), Gul (1991), and Bordalo et al. (2012), may also explain our experimental evidence, since in these theories, outcome spacing of gambles matters (though in different ways). Unlike in PT or CPT where the evaluation of a prospect is

---

[16] The common understanding is that decision weights are a function of probability weights, which reflect an intrinsic and perhaps context-free individual attitude towards chances.

independent of the alternatives available, a common feature of most of these theories is that they have modified the classic dual structure to take into account the impact of emotions or other psychological traits arising from comparisons of risky alternatives. This would allow the 'decision weighting component' of decision making to interact with the contextual factors of decision making in a more psychologically intuitive way. Therefore, in light of our findings, we think the modeling approach that takes into account the comparative and context-dependent nature of decision making is a more promising one, provided that not too much parsimony or interpretability is traded off.

Appendices for Chapter 4

Appendix 4.1 Details of the lotteries used in our experiment

In Table 4.7 we provide basic statistics for each set of lotteries, with respect to their relative outcome spacing (i.e. ratio *x/y*) and outcome level (i.e. expected payoff).

Table 4.7a: Lottery statistics, *p*=0.09

	Ratio *x/y*				**Expected payoff (£)**			
	mean	median	min	max	mean	median	min	max
HS	8.59	5.45	3.36	27.00	9.51	9.50	2.00	16.98
LS	1.36	1.37	1.12	1.68	9.50	9.52	2.02	16.99
HP	4.92	2.50	1.13	25.78	13.50	13.51	10.01	16.99
LP	5.04	2.54	1.12	27.00	5.51	5.50	2.00	9.00

Table 4.7b: Lottery statistics, *p*=0.21

	Ratio *x/y*				**Expected payoff (£)**			
	mean	median	min	max	mean	median	min	max
HS	8.50	5.41	3.34	25.88	12.50	12.51	4.98	20.00
LS	1.36	1.37	1.11	1.66	12.50	12.48	5.01	19.98
HP	4.93	2.48	1.12	25.67	16.50	16.52	12.98	20.00
LP	4.92	2.52	1.11	25.87	8.50	8.50	4.98	12.03

Table 4.7c: Lottery statistics, *p*=0.33

	Ratio *x/y*				**Expected payoff (£)**			
	mean	median	min	max	mean	median	min	max
HS	8.67	5.51	3.38	28.25	15.51	15.53	7.99	23.03
LS	1.36	1.37	1.12	1.65	15.50	15.48	8.02	22.99
HP	4.95	2.51	1.12	26.53	19.51	19.48	16.01	23.03
LP	5.09	2.52	1.12	28.25	11.50	11.50	7.99	15.03

Table 4.8 below shows a list of all the 96 lotteries for our experiment.

Table 4.8: The list of all lotteries for our experiment ((x, p; y, 1-p) with $x>y$)

	p=0.09		*p*=0.21		*p*=0.33	
	x	*y*	*x*	*y*	*x*	*y*
	£79.90	£3.10	£53.90	£2.10	£45.10	£1.70
	£65.70	£5.60	£50.80	£4.20	£43.70	£3.90
	£58.30	£7.40	£48.50	£6.10	£43.60	£5.40
Set 1	£53.80	£9.00	£46.80	£7.80	£42.90	£7.20
	£51.60	£10.30	£46.40	£9.20	£43.30	£8.50
	£49.40	£11.60	£45.50	£10.70	£43.20	£10.10
	£48.50	£12.80	£45.10	£12.10	£43.40	£11.50
	£47.10	£14.00	£44.80	£13.40	£43.60	£12.90
	£16.20	£0.60	£20.70	£0.80	£22.60	£0.80
	£18.20	£1.50	£21.50	£1.90	£23.30	£2.00
	£20.10	£2.40	£23.10	£2.70	£24.40	£2.90
Set 2	£20.60	£3.50	£23.10	£4.00	£24.90	£4.20
	£21.70	£4.50	£24.30	£4.90	£25.70	£5.20
	£23.30	£5.40	£25.40	£5.90	£26.70	£6.20
	£24.00	£6.40	£26.10	£7.00	£27.30	£7.40
	£25.10	£7.40	£27.20	£8.00	£28.50	£8.40
	£15.50	£9.50	£18.60	£11.50	£21.70	£13.20
	£16.30	£10.50	£19.50	£12.50	£22.30	£14.40
	£16.70	£11.50	£20.00	£13.70	£22.90	£15.60
Set 3	£17.50	£12.60	£20.60	£14.80	£23.40	£16.80
	£17.80	£13.60	£20.90	£16.00	£23.50	£18.30
	£18.20	£14.70	£21.20	£17.10	£24.10	£19.50
	£18.70	£15.70	£21.70	£18.30	£24.60	£20.70
	£18.90	£16.80	£21.80	£19.50	£24.80	£22.10
	£3.20	£1.90	£7.30	£4.40	£10.90	£6.60
	£4.40	£2.90	£8.20	£5.40	£11.60	£7.70
	£5.60	£3.80	£9.30	£6.40	£12.70	£8.70
Set 4	£6.70	£4.80	£10.30	£7.40	£13.70	£9.70
	£7.80	£5.80	£11.30	£8.40	£14.50	£10.80
	£8.40	£6.90	£11.70	£9.60	£14.70	£12.10
	£9.30	£7.90	£12.40	£10.60	£15.60	£13.20
	£10.00	£8.90	£13.00	£11.70	£16.10	£14.40

Appendix 4.2 More details of data analysis

In this part we first show details of our Maximum-Likelihood estimation model, and then how we conduct the Likelihood-Ratio test for the aggregate estimation results. Our estimation model is based on equation (4.2), (4.3), and our error specification that the error term ϵ is normally distributed with zero mean and a variance of $(x-y)\sigma^2$. Hence we can set up a log-likelihood function for ϵ:

$$lnF = -\frac{1}{2}ln(2\pi) - \ln\left(\sigma\sqrt{x-y}\right) - \frac{\epsilon^2}{2(x-y)\sigma^2}$$

Note that x and y here denote lottery outcomes. According to equation (4.2) and (4.3),

$$\epsilon = ce - ((1-w)x^\alpha + wy^\alpha)^{1/\alpha}$$

our final log-likelihood function is

$$lnF = -\frac{1}{2}ln(2\pi) - \ln\left(\sigma\sqrt{x-y}\right) - \frac{(ce - ((1-w)x^\alpha + wy^\alpha)^{\frac{1}{\alpha}})^2}{2(x-y)\sigma^2}$$

We then estimate $\alpha, w,$ and σ to maximize the value of this log-likelihood function, with *x, y* being lottery payoffs and *ce* being subjects' input. We ran this estimation for different sets of lotteries at both aggregate and individual levels. For the latter we ran estimations for each individual. Since at the aggregate level, we cannot perform tests of parameter distributions as we did for individual level results, another test strategy is adopted: the likelihood-ratio test.

To run the likelihood-ratio test for our aggregate level results, we introduced two dummy variables to our model instead of estimating the parameters separately with HS, LS, HP and LP lotteries. Dummy variable HIGHS has value 1 if a lottery belongs to the HS set and value 0 if belongs to the LS set. Dummy variable HIGHP has value 1 if a lottery belongs to the HP set and value 0 if belongs to the LP set. Each one of the model parameters we want to estimate, $\beta \in \{\alpha, w, \sigma\}$, is assumed to depend linearly on the dummy variables in the following way:

$$\beta = \beta_{LR} + \beta_{HR}HIGHR \quad \text{for the test of outcome-spacing effect}$$

$$\beta = \beta_{LP} + \beta_{HP}HIGHP \quad \text{for the test of outcome-level effect}$$

There is an outcome-spacing (outcome level) effect if the estimated β_{HR} (β_{HP}) is significantly different from zero. With this setup, we can use the likelihood-ratio test to test the null hypothesis that β_{HR} or β_{HP} equals zero. It is worth mention that estimating the model with the dummy setup specified above yields exactly the same estimation results as what we reported in Table 4.2, which were obtained by estimating the parameters separately with each of the four different sets of lotteries (HS, LS, HP, LP).

Appendix 4.3 More details about estimation results

Appendix 4.3.1 Statistical tests of aggregate-level estimates

Table 4.9 shows details of the Likelihood-Ratio test results presented in Table 4.2, including the Likelihood-Ratio Chi-square value (LR chi2) and the corresponding p-value for each pair of parameter estimates and for each probability. Those denoted with ** indicate the corresponding null hypothesis is rejected at the level of 5%.

Table 4.9: Likelihood-Ratio test results for Table 4.2

Null Hypothesis	*p*=0.09		*p*=0.21		*p*=0.33	
	LR chi2	p-value	LR chi2	p-value	LR chi2	p-value
$\alpha_{HS} = \alpha_{LS}$	0.67	0.412	0.47	0.494	10.48	0.001**
$w_{HS} = w_{LS}$	0.83	0.362	5.00	0.025**	7.76	0.005**
$\sigma_{HS} = \sigma_{LS}$	1875.71	0.000**	1286.12	0.000**	1141.12	0.000**
$\alpha_{HP} = \alpha_{LP}$	12.38	0.000**	0.00	0.969	0.21	0.649
$w_{HP} = w_{LP}$	3.56	0.060	1.17	0.279	0.01	0.917
$\sigma_{HP} = \sigma_{LP}$	325.66	0.000**	159.24	0.000**	157.44	0.000**

Appendix 4.3.2 Statistics of the baseline estimations and robustness check results

Table 4.10, 4.11, 4.12 and 4.13 show respectively the individual-level parameter estimates statistics, including the mean, median, and standard deviations of the distributions over the 84 subjects, for the four estimation setups we use.

Baseline Estimation

Table 4.10a: Statistics of parameters of the baseline estimations (outcome-spacing effect)

Parameter	p=0.09			p=0.21			p=0.33		
	mean	median	s.d.	mean	median	s.d.	mean	median	s.d.
$\widetilde{\alpha}_{HS}$	3.28	1.16	3.42	4.93	1.31	8.38	4.39	1.46	8.05
$\widetilde{\alpha}_{LS}$	3.04	1.59	5.29	1.73	1.21	3.03	-0.68	0.81	3.14
$\widetilde{w}_{HS}$	0.18	0.13	0.19	0.2	0.17	0.17	0.24	0.24	0.2
$\widetilde{w}_{LS}$	0.27	0.23	0.21	0.34	0.31	0.18	0.44	0.43	0.2
$\widetilde{\sigma}_{HS}$	0.71	0.57	0.47	0.66	0.58	0.45	0.71	0.63	0.42
$\widetilde{\sigma}_{LS}$	0.27	0.25	0.13	0.29	0.27	0.13	0.31	0.29	0.14

Table 4.10b: Statistics of parameters of the baseline estimations (outcome-level effect)

Parameter	p=0.09			p=0.21			p=0.33		
	mean	median	s.d.	mean	median	s.d.	mean	median	s.d.
$\widetilde{\alpha}_{HP}$	0.86	0.7	1.46	0.94	1.01	0.88	1.26	1.13	1.19
$\widetilde{\alpha}_{LP}$	1.76	1.36	1.44	0.95	0.86	0.91	1.31	1.11	0.9
$\widetilde{w}_{HP}$	0.29	0.25	0.18	0.33	0.32	0.18	0.35	0.34	0.17
$\widetilde{w}_{LP}$	0.24	0.19	0.18	0.37	0.37	0.17	0.35	0.32	0.17
$\widetilde{\sigma}_{HP}$	0.6	0.48	0.42	0.53	0.47	0.35	0.62	0.54	0.4
$\widetilde{\sigma}_{LP}$	0.39	0.34	0.2	0.43	0.39	0.21	0.46	0.4	0.21

Robustness Check 1

Table 4.11a: Statistics of parameters of Robustness Check 1 (outcome-spacing effect)

Parameter	p=0.09			p=0.21			p=0.33		
	mean	median	s.d.	mean	median	s.d.	mean	median	s.d.
$\widetilde{\alpha}_{HS}$	3.28	1.16	3.42	4.93	1.31	8.38	4.39	1.46	8.05
$\widetilde{w}_{HS}$	0.18	0.13	0.19	0.2	0.17	0.17	0.24	0.24	0.2
$\widetilde{w}_{LS}$	0.26	0.21	0.18	0.28	0.28	0.17	0.3	0.31	0.18
$\widetilde{\sigma}_{HS}$	0.71	0.57	0.47	0.67	0.58	0.45	0.71	0.63	0.42
$\widetilde{\sigma}_{LS}$	0.29	0.28	0.14	0.31	0.29	0.15	0.35	0.33	0.16

Table 4.11b: Statistics of parameters of Robustness Check 1 (outcome-level effect)

Parameter	*p*=0.09			*p*=0.21			*p*=0.33		
	mean	median	s.d.	mean	median	s.d.	mean	median	s.d.
$\widetilde{\alpha}_{HP}$	0.86	0.7	1.46	0.94	1.01	0.88	1.26	1.13	1.19
$\widetilde{w}_{HP}$	0.29	0.25	0.18	0.33	0.32	0.17	0.35	0.34	0.17
$\widetilde{w}_{LP}$	0.33	0.31	0.2	0.37	0.36	0.18	0.37	0.37	0.18
$\widetilde{\sigma}_{HP}$	0.6	0.48	0.42	0.53	0.47	0.35	0.62	0.54	0.4
$\widetilde{\sigma}_{LP}$	0.44	0.4	0.21	0.46	0.42	0.22	0.49	0.46	0.22

Robustness Check 2

Table 4.12a: Statistics of parameters of Robustness Check 2 (outcome-spacing effect)

Parameter	*p*=0.09			*p*=0.21			*p*=0.33		
	mean	median	s.d.	mean	median	s.d.	mean	median	s.d.
$\widetilde{\alpha}_{LS}$	3.04	1.59	5.29	1.73	1.21	3.03	-0.68	0.81	3.14
$\widetilde{w}_{HS}$	0.27	0.09	0.35	0.32	0.23	0.31	0.54	0.42	0.35
$\widetilde{w}_{LS}$	0.27	0.22	0.21	0.33	0.31	0.18	0.6	0.37	0.19
$\widetilde{\sigma}_{HS}$	0.9	0.74	0.54	0.75	0.67	0.44	0.96	0.78	0.55
$\widetilde{\sigma}_{LS}$	0.27	0.25	0.13	0.28	0.27	0.13	0.31	0.29	0.14

Table 4.12b: Statistics of parameters of Robustness Check 2 (outcome-level effect)

Parameter	*p*=0.09			*p*=0.21			*p*=0.33		
	mean	median	s.d.	mean	median	s.d.	mean	median	s.d.
$\widetilde{\alpha}_{LP}$	1.76	1.36	1.44	0.95	0.86	0.91	1.31	1.11	0.9
$\widetilde{w}_{HP}$	0.21	0.16	0.18	0.34	0.32	0.2	0.34	0.31	0.18
$\widetilde{w}_{LP}$	0.24	0.19	0.18	0.37	0.37	0.17	0.35	0.32	0.17
$\widetilde{\sigma}_{HP}$	0.68	0.54	0.43	0.58	0.5	0.36	0.65	0.6	0.38
$\widetilde{\sigma}_{LP}$	0.39	0.34	0.2	0.43	0.39	0.21	0.45	0.4	0.21

Robustness Check 3

Table 4.13a: Statistics of parameters of Robustness Check 3 (outcome-spacing effect)

Parameter	*p*=0.09			*p*=0.21			*p*=0.33		
	mean	median	s.d.	mean	median	s.d.	mean	median	s.d.
$\widetilde{\alpha}_{ALL}$	1.06	0.83	1.12	0.98	0.96	0.94	1.25	1.18	0.96
$\widetilde{w}_{HS}$	0.26	0.23	0.16	0.33	0.32	0.17	0.35	0.33	0.16
$\widetilde{w}_{LS}$	0.31	0.29	0.18	0.38	0.37	0.17	0.37	0.34	0.16
$\widetilde{\sigma}_{HS}$	0.73	0.63	0.47	0.68	0.59	0.42	0.72	0.64	0.41
$\widetilde{\sigma}_{LS}$	0.29	0.28	0.14	0.3	0.28	0.14	0.33	0.31	0.14

Table 4.13b: Statistics of parameters of Robustness Check 3 (outcome-level effect)

Parameter	p=0.09			p=0.21			p=0.33		
	mean	median	s.d.	mean	median	s.d.	mean	median	s.d.
$\widetilde{\alpha}_{\mathrm{ALL}}$	1.06	0.83	1.12	0.98	0.96	0.94	1.25	1.18	0.96
$\widetilde{w}_{\mathrm{HP}}$	0.26	0.23	0.17	0.33	0.32	0.17	0.35	0.32	0.16
$\widetilde{w}_{\mathrm{LP}}$	0.31	0.29	0.18	0.38	0.36	0.17	0.36	0.36	0.16
$\widetilde{\sigma}_{\mathrm{HP}}$	0.6	0.48	0.42	0.55	0.47	0.37	0.62	0.57	0.38
$\widetilde{\sigma}_{\mathrm{LP}}$	0.45	0.4	0.28	0.45	0.4	0.27	0.47	0.43	0.21

Appendix 4.3.3 A summary of the individual-level statistical test results for different estimation setups

In Table 4.14, we summarize the statistical tests results based on which we draw our main conclusion of this study. Results concerning the effects of outcome spacing and outcome level on the three parameters are organized for each probability level: Table 14a, 14b, 14c show the results for p=0.09, 0.21, and 0.33 respectively. We report two types of hypothesis tests we did, one for medians (one-sided Wilcoxon sign-test) and one for the whole distributions (two-sided Wilcoxon signed-rank test). The former has the null hypothesis that the medians from the two samples are the same, and the latter has the null hypothesis that the two samples have the same distribution. In the table we report the qualitative results and the supporting p-value in the parentheses below For example, in Table 4.14a, the test of the equality of median $\widetilde{\alpha}_{\mathrm{HS}}$ and median $\widetilde{\alpha}_{\mathrm{LS}}$, shows a p-value of 0.586, based on which we can accept the null hypothesis and therefore we use 'HS=LS' to denote this result. Since we used two-sided test of distributions, we only report whether the p-value supports equality of distributions or no equality.

Table 4.14a: Hypothesis test results of parameter medians and distributions (p=0.09)

Estimation Setup	$\tilde{\alpha}$ median	$\tilde{\alpha}$ distribution	$\tilde{w}$ median	$\tilde{w}$ distribution	$\tilde{\sigma}$ median	$\tilde{\sigma}$ distribution
Baseline	HS=LS, (0.586) HP<LP, (0.000)	HS=LS, (0.530) HP≠LP, (0.001)	HS<LS, (0.001) HP=LP, (0.101)	HS≠LS, (0.001) HP≠LP, (0.022)	HS>LS, (0.000) HP>LP, (0.000)	HS≠LS, (0.000) HP≠LP, (0.000)
Robustness Check 1	N/A	N/A	HS<LS, (0.000) HP<LP, (0.000)	HS≠LS, (0.000) HP≠LP, (0.000)	HS>LS, (0.000) HP>LP, (0.006)	HS≠LS, (0.000) HP≠LP, (0.001)
Robustness Check 2	N/A	N/A	HS<LS, (0.003) HP<LP, (0.000)	HS=LS, (0.356) HP≠LP, (0.000)	HS>LS, (0.000) HP>LP, (0.000)	HS≠LS, (0.000) HP≠LP, (0.001)
Robustness Check 3	α=1 (0.156)	α=1 (0.669)	HS<LS, (0.000) HP<LP, (0.000)	HS≠LS, (0.000) HP≠LP, (0.000)	HS>LS, (0.000) HP>LP, (0.001)	HS≠LS, (0.000) HP≠LP, (0.000)

Table 4.14b: Hypothesis test results of parameter medians and distributions (p=0.21)

Estimation Setup	$\tilde{\alpha}$ median	$\tilde{\alpha}$ distribution	$\tilde{w}$ median	$\tilde{w}$ distribution	$\tilde{\sigma}$ median	$\tilde{\sigma}$ distribution
Baseline	HS=LS, (0.380) HP=LP, (0.380)	HS≠LS, (0.041) HP=LP, (0.418)	HS<LS, (0.000) HP<LP, (0.004)	HS≠LS, (0.000) HP≠LP, (0.005)	HS>LS, (0.000) HP>LP, (0.027)	HS≠LS, (0.000) HP≠LP, (0.000)
Robustness Check 1	N/A	N/A	HS<LS, (0.000) HP<LP, (0.002)	HS≠LS, (0.000) HP≠LP, (0.000)	HS>LS, (0.000) HP=LP, (0.230)	HS≠LS, (0.000) HP≠LP, (0.050)
Robustness Check 2	N/A	N/A	HS<LS, (0.046) HP<LP, (0.001)	HS=LS, (0.297) HP≠LP, (0.000)	HS>LS, (0.000) HP>LP, (0.000)	HS≠LS, (0.000) HP≠LP, (0.000)
Robustness Check 3	α=1 (0.586)	α=1 (0.831)	HS<LS, (0.000) HP<LP, (0.000)	HS≠LS, (0.000) HP≠LP, (0.000)	HS>LS, (0.000) HP>LP, (0.024)	HS≠LS, (0.000) HP≠LP, (0.002)

Table 4.14c: Hypothesis test results of parameter medians and distributions (*p*=0.33)

Estimation Setup	$\widetilde{\alpha}$		$\widetilde{w}$		$\widetilde{\sigma}$	
	median	distribution	median	distribution	median	distribution
Baseline	HS>LS, (0.000) HP=LP, (0.445)	HS≠LS, (0.000) HP=LP, (0.518)	HS<LS, (0.000) HP=LP, (0.744)	HS≠LS, (0.000) HP=LP, (0.975)	HS>LS, (0.000) HP>LP, (0.021)	HS≠LS, (0.000) HP≠LP, (0.000)
Robustness Check 1	N/A	N/A	HS<LS, (0.000) HP=LP, (0.230)	HS≠LS, (0.000) HP=LP, (0.061)	HS>LS, (0.000) HP>LP, (0.011)	HS≠LS, (0.000) HP≠LP, (0.001)
Robustness Check 2	N/A	N/A	HS<LS, (0.002) HP=LP, (0.230)	HS≠LS, (0.000) HP=LP, (0.187)	HS>LS, (0.000) HP>LP, (0.000)	HS≠LS, (0.000) HP≠LP, (0.000)
Robustness Check 3	α>1 (0.011)	α=1 (0.088)	HS<LS, (0.038) HP=LP, (0.230)	HS≠LS, (0.003) HP=LP, (0.114)	HS>LS, (0.000) HP>LP, (0.011)	HS≠LS, (0.000) HP≠LP, (0.000)

Appendix 4.4 Experimental Instruction

Welcome to this experiment on decision making which we expect to last about an hour. You will be paid £2 for participating so long as you complete all of the required tasks. In addition, you may earn more depending on your decisions and upon chance. We will pay you in cash at the end of the experiment.

There are some general rules you must follow:

1. Please put away your mobile phones and do not talk to others at any time during the experiment.

2. You will use your computer to make decisions during the experiment. Please only use the computer in the way that you are asked to. Do not close the software that is running or use the computer for any other purpose.

3. You must not open your envelope until instructed by the experimenter.

4. If you have any questions during the experiment, please raise your hand.

If you disobey these rules, you may be asked to leave without payment.

Lotteries

In this experiment, you will make a series of decisions involving lotteries and certain amounts of money. At the end of the experiment, you might receive a sum of money for sure or you might play a lottery to determine your payoff.

The picture on the right shows a sample lottery which we refer to as Lottery A. Playing this lottery results in one of two possible outcomes: winning £8.00 or £2.00. Chances of each outcome are shown as percentages and indicated by the relative sizes of the pie slices. If you play a lottery, your payoff will be determined by a draw of chip from a bag of chips numbered from 1 to 100. For example, for Lottery A, you win £8.00 if the number drawn is from 1 to 10 and win £2.00 otherwise. So there is a 10% chance to win £8.00 and a 90% chance to win £2.00, as shown in Picture 1.
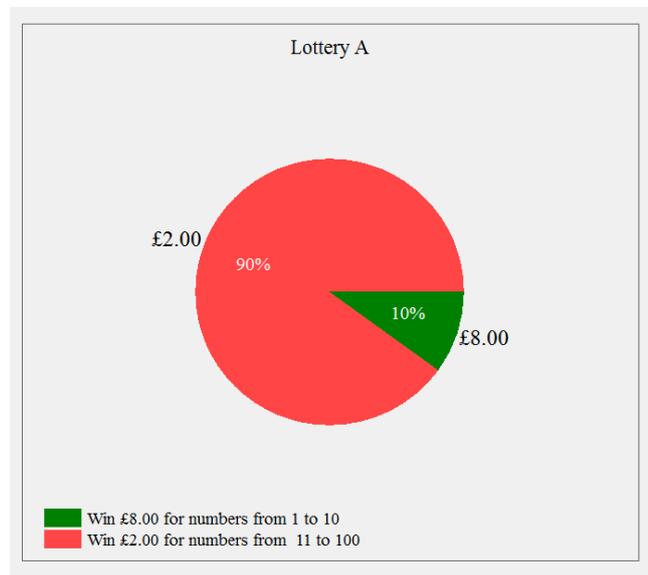
During the experiment, you will see 96 lotteries, labelled from L1 to L96. The lotteries will appear in a random order. Pay attention when a new lottery appears because each lottery is different.

The Task

For each lottery, your task is to enter the amount of money that you think is just as good as the lottery. That is, in a given question you should state the amount of money so that **you are equally willing to play the lottery or instead receive the amount you set for sure.** Besides, we will assume that each of your decision also implies two things: First, you prefer the lottery to any sure amount of money less than the amount

you stated; Secondly, you prefer any sure amount higher than the one you stated to the lottery.

The picture below provides an example of what a task screen will look like. The lottery is shown on the left and you enter the sum that you think is just as good as the lottery on the right.

Please indicate the sure amount of money that you think is just as good as Lottery A

Lottery A

£2.00
90%
10%
£8.00

Win £8.00 for numbers from 1 to 10
Win £2.00 for numbers from 11 to 100

I think playing Lottery A is just as good as receiving £

Click 'Confirm' only if the amount you enter makes you equally willing to play the lottery or receive the amount you set.

Confirm

Picture 2: A sample experimental task

Once you have entered your decision and are ready to proceed, click 'Confirm'. Once you confirm, you won't be able to change your response later.

It is in your best interest to think carefully and, for each lottery, enter the amount that you really think is just as good as it. This is because of the way that we are going to determine your payoff.

How is your payoff determined

To explain how your payoff from the experiment is determined, we need to tell you about what is in your envelope. It contains a slip of paper with two things on it. One is a label identifying one of the 96 lotteries you considered: we will call this your "selected" lottery. For each of the lotteries we have a corresponding envelope, so that

the one you just picked could be any of the 96. The other thing written on the slip in your envelope is an amount of money highlighted in a text box: we will call this your "offer", which can be any amount between the lowest prize and the highest prize of the selected lottery.

To determine your payoff, we will compare the offer in your envelope to the amount that you stated was just as good as the selected lottery:

– If the offer is more than the amount you stated, you will be paid the offer;

 – If the offer is less than the amount you stated, you will play the selected lottery and be paid its outcome;

 – If the offer is equal to the amount you stated, (since you have indicated that you are equally willing to have either) you will be paid the offer.

Notice that you either receive the offer or play the selected lottery. Given that you prefer more money to less, **answering the questions truthfully ensures you get the one you prefer.**

The experiment will start with a small quiz to test your understanding of the instructions. You can only proceed to the main tasks when everyone have completed the quiz successfully.

If you need to ask a question now or later please just raise your hand.

REMEMBER YOU MUST NOT OPEN YOUR ENVELOPE UNTIL INSTRUCTED!

Chapter 5: Conclusions

My goal in this thesis is to contribute to the understanding of individual decision-making under uncertainty, particularly under risk where people choose between lotteries with objectively known probabilities. The three main chapters presented show the major output of my research in the pursuit of the PhD degree.

Chapter 2 proposed a non-fitting method by which the probability weighting function of CPT can be easily approximated, in combination with an existing method that elicits decision weights. The main advantages of this method are its simplicity and efficiency in terms of data requirements. Moreover, no elicitation of the subjective probabilities is needed for measuring probability weighting functions for the case of ambiguity. An experiment of choice under risk was run to allow for the application of the Neo-Lite method and the standard parametric fitting method. Since our method relies on the assumption of the Neo-additive functional form, we also compared the predictive power of this function with typical non-linear functional forms. Our results show that the Neo-Lite method provides a more efficient way, compared to standard parametric fitting, to obtain measurements of the Neo-additive weighting function. The latter performs almost equally well as the non-linear probability weighting functions in terms of in-sample predictions, though the general out-of-sample predictive performance is poor.

This study contributes to the literature mainly in two ways. First, we have proposed an empirical method for the measurement of probability weighting function for CPT, and have shown that this method can be used as a quick and convenient way to obtain some pre-analysis results or to get a sense of the shape of the probability weighting function. Secondly, our experimental findings, that people have inverse-S shaped probability weighting functions for CE tasks but S-shaped

weighting functions for binary choice tasks, add to the literature further evidence of a descriptive limitation of CPT. This then puts into question the extent to which the functions of the CPT model measured with CE tasks can represent individual risk preferences[1].

Chapter 3 is partly motivated by the findings from Chapter 2. In Chapter 3 we propose a model that explains, when preferences are elicited using CE tasks, why individuals are often found to have CPT probability weighting functions that are: (i) non-linear and inverse-S shaped; (ii) different for the domain of gains and losses; (iii) sensitive to the lottery outcomes. In addition, the model also allows the CPT probability weighting function to be considerably different for CE tasks and for binary choices. Our theory is based on the psychologically intuitive mechanism of *attention-based state-weighting,* and takes into account the comparative nature of decision-making (i.e. evaluation of a risky option can depend upon features of the broader set of alternatives in which it is embedded). With Propositions 1, 2 and 3, we have shown that our theory is equivalent to CPT with a much more complicated decision weighting function.

We consider the contribution of Chapter 3 to be two-fold. First, it provides a theoretical explanation as to why people are often found to have inverse-S shaped non-linear probability weighting functions under CPT and why the weighting function can differ between the outcome domain of gains and of losses. The answer to these questions does not necessarily have anything to do with human psychophysics to distort likelihoods (if the latter exists). Instead we have shown that these findings can be accounted for with a decision-making process which does not presume context-independent distortions of objective probabilities. Secondly, our theory provides a coherent accommodation for evidence that could not be explained by EUT (e.g. Kahneman & Tversky 1979) and that could not be explained by CPT (e.g. Birmbaun 2008, Fehr-Duda et al. 2010), including the preference reversal

---

[1] It remains a controversial question, since the finding of preference reversals, whether valuations of lotteries or choices between lotteries better represents individual risk preferences. The argument here is not against the use of the CE method for preference elicitation, but against the combination of the use of CE and the use of the CPT model. For many other theories that can explain preference reversals, such as Loomes & Sugden (1982, 1983), Goldstein & Einhorn (1987), Schmidt et al. (2008), Bordalo et al. (2012), the CE approach may not be a problem.

phenomenon (e.g. Grether & Plott 1979, Cubitt et al. 2004). The two key ingredients that underlie our theory are the psychological mechanism of attention-based state-weighting (Bordalo et al. 2012) and reference-dependent decision making (e.g. Kahneman & Tversky 1979, 1992). Neither ingredients are brand-new ideas, but the combination of the two ingredients generates a surprisingly powerful theory. We think the potential of this theory presented in Chapter 3 is yet to be fully explored.

Complementary to Chapter 3, the study of Chapter 4 pursues two goals. First, we investigate experimentally the effects of contextual factors related to lottery outcomes on decision weighting. Our data suggests that outcome spacing has a significant effect on decision weighting at both aggregate and individual levels and outcome level only has significant effect at individual level. These results can be accommodated by our theory presented in Chapter 3. The second goal is to test a specific prediction derived from Proposition 1 of Chapter 3, regarding the effect of outcome spacing and the dependence of that effect on value function curvatures. We find that this prediction is mostly supported for subjects with convex power value functions, but less supported for subjects with concave value functions.

The main contribution of the last study is to show clear-cut evidence that, decision weights (of the best lottery outcomes) are affected by both lottery outcome spacing and outcome level. Besides, we have isolated the effect of outcome spacing from outcome level, and have shown that in our experiment, people are more sensitive to the former than to the latter[2]. These results challenge the CPT model, especially its rank-dependent decision weighting function, and provide empirical grounds for the idea that decision weights, in the framework of PT and CPT, can depend on contextual factors of the decision problem systematically and significantly.

Admittedly, there are various ways this study could be extended or improved. For example, more research could be done to see the effects of outcomes on decision weighting for lotteries with a high probability to win the best outcome. Moreover,

---

[2] This result, however, may not be general, because our manipulations of differences in lottery outcome levels are weaker than our manipulations of differences in outcome spacing.

robustness of our results could be further checked by using alternative preference elicitation techniques, such as choice-list tasks, rather than the stated-CE tasks. Nevertheless, we think the study of Chapter 4 has generated valuable insights regarding the way we should interpret decision weights. In light of the evidence presented, it is perhaps too narrow to interpret decision weighting as context-independent probability weighting, as in PT. There might exist a part of deviation of decision weights (from objective probabilities) that is individual-unique and context-independent, but it remains a question whether we should build that part into theoretical models (as in CPT), because it is rather difficult to observe the context-independent part of decision weighting even if we know it exists.

Overall, this thesis has investigated several issues related to probability weighting and decision weighting, two key concepts from two of the most influential and widely used models of behavioral economics. The three studies presented in this thesis have updated our understanding of the two concepts and of the limitations of CPT. Standing on evidence from our studies, we are no longer sure the context-independent probability weighting is still a useful concept for descriptive models or for empirical research. Neither are we sure about whether CPT is still the right model through which we can best understand individual behaviors under risk, or whether CPT is the right model to extend. We think it is better for future theories of choice to take into account the comparative nature of decision-making. Although Chapter 3 has already provided such a model, there is certainly scope for further research, both theoretical and empirical, to be done. For example, more experimental tests could be done regarding other implications of our theory, such as the connection between decision weighting and loss aversion. Regularities about reference point in decisions under risk need also be further investigated and systematized. In addition, all three studies in this thesis can be extended to the case of ambiguity. The future of the area of individual behaviors under uncertainty remains exciting and vigorous.

Bibliography

Abdellaoui, M. (2000). Parameter-free elicitation of utility and probability weighting functions. *Management Science*, 46(11), 1497–1512.

Abdellaoui, M., & Wakker, P. P. (2005). The likelihood method for decision under uncertainty. *Theory and Decision*, 58(1), 3-76.

Abdellaoui, M., Bleichrodt, H., & l'Haridon, O. (2008). A tractable method to measure utility and loss aversion under prospect theory. *Journal of Risk and uncertainty*, *36*(3), 245.\

Abdellaoui, M., Bleichrodt, H., & Paraschiv, C. (2007). Loss aversion under prospect theory: A parameter-free measurement. *Management Science*, 53(10), 1659-1674.

Abdellaoui, M., Diecidue, E., & Öncüler, A. (2011b). Risk preferences at different time periods: An experimental investigation. *Management Science*, *57*(5), 975-987.

Abdellaoui, M., Placido, L., Baillon, A., & Wakker, P. P. (2011a). The Rich Domain of Uncertainty: Source Functions and Their Experimental Implementation. *American Economic Review*, *101*(2), 695-723.

Abdellaoui, M., Vossmann, F., & Weber, M. (2005). Choice-based elicitation and decomposition of decision weights for gains and losses under uncertainty. *Management science*, *51*(9), 1384-1399.

Alarié, Y., & Dionne, G. (2001). Lottery decisions and probability weighting function. *Journal of Risk and Uncertainty*, 22, 21–33.

Allais, M. (1953). L'extension des théories de l'équilibre économique général et du rendement social au cas du risque. *Econometrica, Journal of the Econometric Society*, 269-290.

Allais, M. (1988). The General Theory of Random Choices in Relation to the Invariant Cardinal Utility Function and the Specific Probability Function. The (U,

θ) Model A General Overview. In *Risk, decision and rationality* (pp. 231-289). Springer Netherlands.

Beattie, J., & Loomes, G. (1997). The impact of incentives upon risky choice experiments. *Journal of Risk and Uncertainty*, *14*(2), 155-168.

Bell, D. (1985). Disappointment in decision making under uncertainty. *Operations Research, 33*, 1–27.

Binswanger,H. P. (1981).Attitudes toward risk: Theoretical implications of an experiment in rural India. *Economic Journal, 91*, 867–890.

Blavatskyy, P. R. (2005). Back to the St. Petersburg paradox?. *Management Science*, *51*(4), 677-678.

Blavatskyy, P. (2006). Error propagation in the elicitation of utility and probability weighting functions. *Theory and Decision*, 60(2-3), 315-334.

Blavatskyy, P. R. (2007). Stochastic expected utility theory. *Journal of Risk and Uncertainty*, *34*(3), 259-286.

Bleichrodt, H., & Pinto, J. L. (2000). A parameter-free elicitation of the probability weighting function in medical decision analysis. *Management science*, *46*(11), 1485-1496.

Bleichrodt, H., van Rijn, J., & Johannesson, M. (1999). Probability weighting and utility curvature in QALY-based decision making. *Journal of mathematical psychology*, *43*(2), 238-260.

Bombardini, M., & Trebbi, F. (2012). Risk aversion and expected utility theory: an experiment with large and small stakes. *Journal of the European Economic Association*, *10*(6), 1348-1399.

Booij, A. S., Van Praag, B. M., & Van De Kuilen, G. (2010). A parametric analysis of prospect theory's functionals for the general population. *Theory and Decision*, *68*(1-2), 115-148.

Bordalo, P., Gennaioli, N., & Shleifer, A. (2012). Salience theory of choice under risk. *The Quarterly journal of economics*, *127*(3), 1243-1285.

Bouchouicha, R., & Vieider, F. M. (2016). *Accommodating stake effects under prospect theory* (No. em-dp2016-03). Henley Business School, Reading

University.

Brandstätter, E., Kühberger, A., & Schneider, F. (2002). A cognitive‑emotional account of the shape of the probability weighting function. *Journal of Behavioral Decision Making*, *15*(2), 79-100.

Bruhin, A., Fehr‑Duda, H., & Epper, T. (2010). Risk and rationality: Uncovering heterogeneity in probability distortion. *Econometrica*, *78*(4), 1375-1412.

Butler, David, J., and Graham C. Loomes. (2007). Imprecision as an account of the preference reversal phenomenon. *American Economic Review*, 97(1): 277-297.

Camerer, C. F. (1992). Recent tests of generalizations of expected utility theory. *Utility theories: Measurements and applications*, 207-251.

Camerer, C. F., & Ho, T.-H. (1994). Violations of the betweenness axiom and nonlinearity in probability. *Journal of risk and uncertainty*, 8(2), 167-196.

Camerer, C. F., & Hogarth, R. M. (1999). The effects of financial incentives in experiments: A review and capital-labor-production framework. *Journal of Risk and Uncertainty*, 19, 7–42.

Camerer, C., & Weber, M. (1992). Recent developments in modeling preferences: Uncertainty and ambiguity. *Journal of risk and uncertainty,* 5(4), 325-370.

Chateauneuf, A., Eichberger, J., & Grant, S. (2007). Choice under uncertainty with the best and worst in mind: Neo-additive capacities. *Journal of Economic Theory*, *137*(1), 538-567.

Cohen, M. (1992). Security level, potential level, expected utility: A three-criteria decision model under risk. *Theory and Decision*, 33(2), 101-134.

Conte, A., Hey, J. D., & Moffatt, P. G. (2011). Mixture models of choice under risk. *Journal of Econometrics*, 162(1), 79-88.

Cubitt, R. P., Munro, A., & Starmer, C. (2004). Testing explanations of preference reversal. *The Economic Journal*, *114*(497), 709-726.

Dave, C., Eckel, C. C., Johnson, C. A., & Rojas, C. (2010). Eliciting risk preferences: When is simple better?. *Journal of Risk and Uncertainty*, *41*(3), 219-243.

Diecidue, E., & Wakker, P. P. (2001). On the intuition of rank-dependent utility. *Journal of Risk and Uncertainty*, 23, 281–298.

Diecidue, E., Schmidt, U., & Zank, H. (2009). Parametric weighting functions. *Journal of Economic Theory*, *144*(3), 1102-1118.

Diecidue, E., Wakker, P. P., & Zeelenberg, M. (2007). Eliciting decision weights by adapting de finetti's betting-odds method to prospect theory. *Journal of risk and uncertainty*, 34(3), 179-199.

Dyer, J. S., & Sarin, R. K. (1982). Relative risk aversion. *Management Science*, *28*(8), 875-886.

Edwards, W. (1954). The theory of decision making. *Psychological bulletin*, *51*(4), 380.

Edwards, W. (1962). Subjective probabilities inferred from decisions. *Psychological review*, *69*(2), 109.

Ellsberg, D. (1961). Risk, ambiguity, and the Savage axioms. *The quarterly journal of economics*, 643-669.

Etchart-Vincent, N. (2004). Is probability weighting sensitive to the magnitude of consequences? An experimental investigation on losses. *Journal of Risk and Uncertainty*, 28(3), 217–235.

Etchart-Vincent, N. (2009). Probability weighting and the 'level'and 'spacing'of outcomes: An experimental study over losses. *Journal of Risk and Uncertainty*, *39*(1), 45-63.

Fehr-Duda, H., Bruhin, A., Epper, T., & Schubert, R. (2010). Rationality on the rise: Why relative risk aversion increases with stake size. *Journal of Risk and Uncertainty*, 40(2), 147-180.

Fehr-Duda, H., de Gennaro, M., & Schubert, R. (2006). Gender, financial risks, and probability weights. *Theory and Decision*, 60, 283–313.

Fennema, H., & Wakker, P. (1997). Original and cumulative prospect theory: A discussion of empirical differences. *Journal of Behavioral Decision Making*, *10*, 53-64.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental economics*, 10(2), 171-178.

Fox, C. R., & Tversky, A. (1998). A belief-based account of decision under

uncertainty. *Management Science,* 44(7), 879-895.

Gilboa, I. (1988). A combination of expected utility and maxmin decision criteria. *Journal of Mathematical Psychology,* 32(4), 405-420.

Goldstein, W. M., & Einhorn, H. J. (1987). Expression theory and the preference reversal phenomena. *Psychological review*, 94(2), 236.

Gonzalez, R., & Wu, G. (1999). On the shape of the probability weighting function. *Cognitive Psychology*, 38, 129–166.

Greiner, B. Subject pool recruitment procedures: organizing experiments with ORSEE. *Journal of the Economic Science Association*, 1-12.

Grether, D. M., & Plott, C. R. (1979). Economic theory of choice and the preference reversal phenomenon. *The American Economic Review*, 69(4), 623-638.

Gul, F. (1991). A theory of disappointment aversion. *Econometrica, 59*(3), 667–686.

Harbaugh, W. T., Krause, K., & Vesterlund, L. (2002). Risk attitudes of children and adults: Choices over small and large probability gains and losses. *Experimental Economics*, *5*(1), 53-84.

Harless, D. W., & Camerer, C. F. (1994). The predictive utility of generalized expected utility theories. *Econometrica*, 62(6), 1251–1289.

Harrison, G. W. (1992). Theory and misbehavior of first-price auctions: Reply. *The American Economic Review*, *82*(5), 1426-1443.

Harrison, G. W. (2008). Maximum likelihood estimation of utility functions using Stata. University of Central Florida, Working Paper, 06-12.

Harrison, G. W., & Rutström, E. E. (2009). Expected utility theory and prospect theory: One wedding and a decent funeral. *Experimental Economics*, *12*(2), 133-158.

Heath, C., & Tversky, A. (1991). Preference and belief: Ambiguity and competence in choice under uncertainty. *Journal of risk and uncertainty,* 4(1), 5-28.

Hey, J. D. (1995). Experimental investigations of errors in decision making under risk. *European Economic Review*, *39*(3), 633-640.

Hey, J. D., & Orme, C. (1994). Investigating generalizations of expected utility theory using experimental data. *Econometrica, 62*(6), 1291–1326.

Hey, J. D., Lotito, G., & Maffioletti, A. (2010). The descriptive and predictive adequacy of theories of decision making under uncertainty/ambiguity. *Journal of risk and uncertainty*, 41(2), 81-111.

Humphrey, S. J., & Verschoor, A. (2004). The probability weighting function: Experimental evidence from Uganda, India and Ethiopia. *Economics Letters*, 84, 419–425.

Kachelmeier, S. J., & Shehata, M. (1992). Examining risk preferences under high monetary incentives: Experimental evidence from the People's Republic of China. *American Economic Review*, 82(5), 1120–1141.

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

Kahneman, D., & Tversky, A. (1984). Choices, values, and frames. *American psychologist*, *39*(4), 341.

Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1991). Anomalies: The endowment effect, loss aversion, and status quo bias. *The Journal of Economic Perspectives*, 193-206.

Keynes, J. M. (1921). A treatise on probability: Courier Dover Publications.

Kilka, M., & Weber, M. (2001). What determines the shape of the probability weighting function under uncertainty?. *Management science*, *47*(12), 1712-1726.

Knight, F. H. (1921). Risk, uncertainty and profit. New York: Hart, Schaffner and Marx.

Kontek, K. (2016). A critical note on Salience Theory of choice under risk. *Economics Letters*, *149*, 168-171.

Kőszegi, B., & Rabin, M. (2006). A model of reference-dependent preferences. *The Quarterly Journal of Economics*, 1133-1165.

Kothiyal, A., Spinu, V., & Wakker, P. P. (2014). An experimental test of prospect theory for predicting choice under ambiguity. *Journal of risk and uncertainty*, 48(1), 1-17.

Kuehberger, A., Schulte-Mecklenbeck, M., & Perner, J. (1999). The effects of framing, reflection, probability, and payoff on risk preference in choice tasks.

Organizational Behavior and Human Decision Processes, 78(3), 204–231.

Kusev, P., van Schaik, P., Ayton, P., Dent, J., & Chater, N. (2009). Exaggerated risk: Prospect theory and probability weighting in risky choice. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35(6), 1487.

Lattimore, P. K., Baker, J. R., & Witte, A. D. (1992). The influence of probability on risky choice: a parametric examination. *Journal of Economic Behavior and Organization*, 17, 377–400.

Lichtenstein, S., & Slovic, P. (1971). Reversals of preference between bids and choices in gambling decisions. *Journal of experimental psychology*, 89(1), 46.

Loomes, G. (2005). Modelling the stochastic component of behaviour in experiments: Some issues for the interpretation of data. *Experimental Economics*, *8*(4), 301-323.

Loomes, G., & Sugden, R. (1982). Regret theory: An alternative theory of rational choice under uncertainty. *The economic journal*, *92*(368), 805-824.

Loomes, G., & Sugden, R. (1983). A rationale for preference reversal. *The American Economic Review*, 428-432.

Loomes, G., & Sugden, R. (1986). Disappointment and dynamic consistency in choice under uncertainty. *The Review of Economic Studies*, *53*(2), 271-282.

Loomes, G., Moffatt, P. G., & Sugden, R. (2002). A microeconometric test of alternative stochastic theories of risky choice. *Journal of risk and uncertainty*, 24(2), 103-130.

Loomes, G., Starmer, C., & Sugden, R. (1989). Preference reversal: information-processing effect or rational non-transitive choice? *The Economic Journal*, 140-151.

Luce, R. D., & Fishburn, P. C. (1991). Rank-and sign-dependent linear utility models for finite first-order gambles. *Journal of risk and Uncertainty*, *4*(1), 29-59.

Prelec, D. (1998). The probability weighting function. *Econometrica*, 66(3), 497–527.

Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior & Organization*, *3*(4), 323-343.

Quiggin, J. (2012). *Generalized expected utility theory: The rank-dependent model*.

Organizational Behavior and Human Decision Processes, 78(3), 204–231.

Kusev, P., van Schaik, P., Ayton, P., Dent, J., & Chater, N. (2009). Exaggerated risk: Prospect theory and probability weighting in risky choice. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35(6), 1487.

Lattimore, P. K., Baker, J. R., & Witte, A. D. (1992). The influence of probability on risky choice: a parametric examination. *Journal of Economic Behavior and Organization*, 17, 377–400.

Lichtenstein, S., & Slovic, P. (1971). Reversals of preference between bids and choices in gambling decisions. *Journal of experimental psychology*, 89(1), 46.

Loomes, G. (2005). Modelling the stochastic component of behaviour in experiments: Some issues for the interpretation of data. *Experimental Economics*, *8*(4), 301-323.

Loomes, G., & Sugden, R. (1982). Regret theory: An alternative theory of rational choice under uncertainty. *The economic journal*, *92*(368), 805-824.

Loomes, G., & Sugden, R. (1983). A rationale for preference reversal. *The American Economic Review*, 428-432.

Loomes, G., & Sugden, R. (1986). Disappointment and dynamic consistency in choice under uncertainty. *The Review of Economic Studies*, *53*(2), 271-282.

Loomes, G., Moffatt, P. G., & Sugden, R. (2002). A microeconometric test of alternative stochastic theories of risky choice. *Journal of risk and uncertainty*, 24(2), 103-130.

Loomes, G., Starmer, C., & Sugden, R. (1989). Preference reversal: information-processing effect or rational non-transitive choice? *The Economic Journal*, 140-151.

Luce, R. D., & Fishburn, P. C. (1991). Rank-and sign-dependent linear utility models for finite first-order gambles. *Journal of risk and Uncertainty*, *4*(1), 29-59.

Prelec, D. (1998). The probability weighting function. *Econometrica*, 66(3), 497–527.

Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior & Organization*, *3*(4), 323-343.

Quiggin, J. (2012). *Generalized expected utility theory: The rank-dependent model*.

Springer Science & Business Media.

Rottenstreich, Y., & Hsee, C. K. (2001). Money, kisses, and electric shocks: On the affective psychology of risk. *Psychological science*, *12*(3), 185-190.

Saha, A. (1993). Expo-power utility: A 'flexible'form for absolute and relative risk aversion. *American Journal of Agricultural Economics*, *75*(4), 905-913.

Savage, L. J. (1954). *The foundations of statistics*: Courier Dover Publications.

Savage, L. J. (1972). *The foundations of statistics*. Courier Corporation.

Schmeidler, D. (1989). Subjective probability and expected utility without additivity. *Econometrica: Journal of the Econometric Society*, 571-587.

Schmidt, U., Starmer, C., & Sugden, R. (2008). Third-generation prospect theory. *Journal of risk and uncertainty*, 36(3), 203-223.

Scholten, M., & Read, D. (2014). Prospect theory and the "forgotten" fourfold pattern of risk preferences. *Journal of Risk and Uncertainty*, *48*(1), 67-83.

Slovic, P., Lichtenstein, S., & Fischhoff, B. (1988). *Decision making*. Wiley.

Starmer, C. (2000). Developments in non-expected utility theory: The hunt for a descriptive theory of choice under risk. *Journal of economic literature*, *38*(2), 332-382.

Starmer, C., & Sugden, R. (1989). Violations of the independence axion in common ratio problems: An experimental test of some competing hypotheses. *Annals of Operations Research*, *19*(1), 79-102.

Stott, H. P. (2006). Cumulative prospect theory's functional menagerie. *Journal of risk and uncertainty*, 32(2), 101-130.

Teitelbaum, J. C. (2007). A unilateral accident model under ambiguity. *The Journal of Legal Studies*, 36(2), 431-477.

Trepel, C., Fox, C. R., & Poldrack, R. A. (2005). Prospect theory on the brain? Toward a cognitive neuroscience of decision under risk. *Cognitive brain research*, *23*(1), 34-50.

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological review*, *90*(4), 293.

Tversky, A., & Kahneman, D. (1985). The framing of decisions and the psychology

of choice. In *Environmental Impact assessment, technology assessment, and risk analysis* (pp. 107-129). Springer, Berlin, Heidelberg.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and uncertainty*, *5*(4), 297-323.

Tversky, A., & Wakker, P. (1995). Risk attitudes and decision weights. *Econometrica: Journal of the Econometric Society*, 1255-1280.

van de Kuilen, G., & Wakker, P. P. (2011). The midweight method to measure attitudes toward risk and ambiguity. *Management Science*, 57(3), 582-598.

Vieider, F. M., Cingl, L., Martinsson, P., & Stojic, H. (2013). *Separating attitudes towards money from attitudes towards probabilities: Stake effects and ambiguity as a test for prospect theory* (No. SP II 2013-401). WZB Discussion Paper.

Viscusi, W. K., & Evans, W. N. (2006). Behavioral probabilities. *Journal of Risk and Uncertainty*, *32*(1), 5-15.

Von Neumann, J., & Morgenstern, O. (1945). *Theory of games and economic behavior*. Princeton, NJ: Princeton University Press.

Wakker, P. (1989). Transforming probabilities without violating stochastic dominance. In *Mathematical psychology in progress* (pp. 29-47). Springer Berlin Heidelberg.

Wakker, P. P. (2008). Explaining the characteristics of the power (CRRA) utility function. *Health Economics*, 17, 1329–1344

Wakker, P. P. (2010). *Prospect theory: For risk and ambiguity*. Cambridge university press.

Wakker, P., & Deneffe, D. (1996). Eliciting von Neumann-Morgenstern utilities when probabilities are distorted or unknown. *Management science*, *42*(8), 1131-1150.

Wu, G., & Gonzalez, R. (1996). Curvature of the probability weighting function. *Management science*, *42*(12), 1676-1690.

Wu, G., Zhang, J., & Abdellaoui, M. (2005). Testing prospect theories using probability tradeoff consistency. *Journal of risk and uncertainty*, 30(2), 107-131.