Investigation of a Behavioural Model for Financial Decision Making

Courtney Alyssandra Pitcher

Magdalen College

University of Oxford

A dissertation submitted for the degree of

MSc Mathematical & Computational Finance

2008

Abstract

Many economic models assume that individuals make decisions by maximizing their expected utility. Expected utility theory was developed to explain the way people behave when faced with choices under risk and uncertainty. However, the explanatory power of this theory has come into question because of systematic violations that have been observed in practice. This paper summarizes these violations and analyses a new theoretical framework that was introduced to overcome these violations called prospect theory. This theory was first proposed by Kahneman and Tversky in 1979, but the theory was later modified to become cumulative prospect theory. The purpose of this paper is to examine this new theory and to apply its framework to the lottery market. The parameters of the functional form of cumulative prospect theory are estimated. A value function with rapidly diminishing sensitivity, and a decision weighting function that was essentially a step function is found. The implications of these results are examined, and these results are compared to estimates given in the literature.

Contents

Chapter 1

Background

1.1 Introduction

Recently, behaviourial economics has been making its way to the forefront in the study of economics. Behavioural economics incorporates psychology into economics and uses it to explain various things throughout the economy [10]. For example, the financial market is made up of many individuals making decisions under risk and uncertainty. In order to understand how the market works one must understand how and why these individuals behave the way they do. This becomes particularly important when trying to predict outcomes in the market. Traditional economic models assume that individuals make decisions by maximizing their expected utility. Expected utility theory is the classical model that is viewed by many as a normative model, which provides an account of the way rational individuals behave when faced with choices under risk and uncertainty. The assumption that people are rational agents, according to expected utility theory, is the same thing as saying that they are exempt from errors in judgment, display self-control, and are averse to risk; however, normal individuals do not always behave in this way and the markets are the aggregate of these normal individuals. [10] Systematic violations of expected utility theory are seen in many different situations, and have been thoroughly documented. This called for a new theory of decision making under risk and uncertainty. Kahneman and Tversky [17] [31] have proposed the most tractable theory so far. What was originally called prospect theory was later modified to become cumulative prospect theory.

The purpose of this paper is to use cumulative prospect theory to build a behavioural model that captures the way individuals behave when faced with a decision under risk and uncertainty that then can be applied to areas of financial decision making, for example, asset allocation, asset pricing, and portfolio optimization. We use the parametric forms proposed by Kahneman and Tversky [31] and we estimate the parameters using data from the lottery market to construct a collective model for a single representative agent. We analyze these estimates and their application to out-of-sample situations. We also compare them to others that have been estimated in the literature.

The paper is organized as follows: we will first overview expected utility theory and its behavioural violations, prospect theory will be introduced followed by an analysis of cumulative prospect theory. The parametric form of cumulative prospect theory will be examined and then the parameters will be estimated. Investigations of these parameter estimates will be discussed, and then finally further applications of cumulative prospect theory will be introduced.

1.2 Expected utility theory

The theory of utility is used based on assumption that individuals evaluate money in terms of the utility they gain from it. If an individual is faced with a choice utility theory can be used to explain and possibly predict the choice a rational individual would or should choose. A *prospect* or a gamble is a finite set of outcomes assigning each outcome a probability of occurrence which is either known or unknown to the decision maker beforehand. [12] An outcome, x_i, is defined as the monetary result of an event, A_i. We will assume that an event can have only one outcome, and for the purposes of this paper, all outcomes are defined in terms of monetary value. It will also be assumed, for now, that the probability of event A_i happening, with a monetary outcome x_i, is known and is given by p_i. We denote a prospect by $(p_1, x_1; \ldots; p_n, x_n)$, and for convenience this will be shortened to $(\mathbf{x}, \mathbf{p})$, where $\mathbf{x} = (x_1, \ldots, x_n)$ is the set of outcomes of the prospect and $\mathbf{p} = (p_1, \ldots, p_n)$ is the corresponding probability associated with the each outcome. For simplicity we do not include outcomes with a zero probability of occurring. The following equation defines what is meant by the expected utility of a prospect $(\mathbf{x}, \mathbf{p})$.

$$E[u(\mathbf{x})] = V(\mathbf{x}, \mathbf{p}) = p_1 u(x_1) + p_2 u(x_2) + \ldots + p_n u(x_n) \tag{1.2.1}$$

$$= \sum_{i=1}^{n} p_i u(x_i) \tag{1.2.2}$$

where the utility of x_i is given by an individuals utility function for money, $u(x)$ and the expected utility of a prospect is denoted by $V(\mathbf{x}, \mathbf{p})$. This paper assumes that individuals make decisions with a finite set of outcomes; however, this theory and all other further theories can be extended to a continuous set of outcomes [31].

The motivation behind this theory is that individuals, instead of making decisions by maximizing the expected value of prospects, they maximize their expected utility. One of the main assumptions of expected utility theory is that people are risk averse meaning that the marginal utility of a unit increase of monetary gain diminishes as an individual becomes more wealthy. This notion is quite intuitive; the more you have of something the less you appreciate it. This implies that the utility function is concave, i.e. $u''(x) < 0$. Risk aversion means that individuals prefer having for certain the expected value of the prospect, $(\mathbf{x}, \mathbf{p})$, than the risky prospect itself. In other words,

$$V(\mathbf{x}, \mathbf{p}) \leq u(E[\mathbf{x}]) = u(p_1 x_1 + \ldots + p_n x_n). \tag{1.2.3}$$

For example, most people would prefer \$500 than a $50 - 50$ chance of \$1000 or nothing. When deciding between two propspect, $(\mathbf{x}, \mathbf{p})$ and $(\mathbf{y}, \mathbf{q})$ an individual will choose $(\mathbf{x}, \mathbf{p})$ if the expected utility is greater, i.e. if $V(\mathbf{x}, \mathbf{p}) > V(\mathbf{y}, \mathbf{q})$.

Expected utility theory is a classsical theory of decision making that combines the principle of mathematical expectation with the assumption of risk aversion. However, this theory does not always hold true because individuals systematically violate the basic premises [32]. The next section gives an overview of these violations.

1.2.1 Violations of expected utility theory

The consistent discrepancies of expected utility theory can be summarized into four main categories of behaviour: loss aversion, nonlinear preferences, risk seeking, and source dependence. These violations or analomies will be explained below.

1. **Loss Aversion**

    - This refers to the fact that losses are felt for longer and have more impact on individuals than gains do. For example, losing your house would be a much more significant event than acquiring a new house. Another example is the reluctance to accept losses on the stock market: the volume of trades tends to be higher when the prices are rising than when they are falling. [30]

    This kind of rationale makes sense from a psychological standpoint because individuals' perception is attuned to differences rather than to absolute magnitudes. For example, brightness, loudness, and temperature. An individual's past and present experiences define their reference point and they view the world in relation to this reference point. This same principle can be applied to wealth. [17] The classical expected utility theory cannot take this kind of behaviour into account because it looks at utility in terms of final assets- not gains and losses. There is also extensive experimental evidence that losses have greater impact on preferences than gains [30].

Loss aversion is one factor that motivates redefining the theory in terms of gains and losses.

2. **Risk seeking**

   - Individuals are not always risk averse, sometimes they are risk seeking. For example, gambling on unfair prospects such as the lottery.

   Kahneman and Tversky [31] find evidence that there exists a fourfold risk pattern structure; individuals are risk seeking for losses and risk averse for gains for prospects with moderate to high probabilities, and individuals are risk averse for losses and risk seeking for gains for prospects with low probabilities. So individuals tend to prefer a substantial probability of a large loss than a sure smaller loss. This is consistent with why people will hold onto stock losers longer than they should, and why individuals buy insurance.

   Friedman and Savage [14], and [20] tried resolve the combination of risk seeking and risk aversion in terms of a utility function with both concave and convex regions. The problem with this is that the fourfold pattern exists over a wide range of payoffs and cannot be explained by just a utility function. This suggests a nonlinear transformation of the probability scale [32]. Also, risk seeking in the domain of losses gives another reason to define the model in terms of gains and losses.

3. **Nonlinear preferences**

   - Preferences between risky prospects are not linear in the probabilities, the difference between a probability of 0.99 and 1 has more impact on preferences than the difference between 0.1 and 0.11.

   A classic example is to consider Russian roulette. Four of the six chambers of a gun have a bullet in them. They spin the chamber and shoot the gun pointed at their head. Individuals will pay more to decrease the number of bullets from 1 to 0 than from 4 to 3, but according to expected utility theory this should not be the case. The reason is because people weigh probabilities nonlinearly. They value the reduction of the probability of being shot from $1/6$ to 0 more than $4/6$ to $3/6$ [22]. It has also been found that individuals overweight small probabilities and underweight moderate to large probabilities. This implies a response to probabilities in a nonlinear manner, which again suggests a nonlinear transformation of the probability scale.

4. **Source dependence**

- An individual's preferences depend not only on the degree of uncertainty but also the on the source of uncertainty.

Individuals prefer to bet on their own judgment and in their own area of competence. They will even pay a premium for this. Individuals will also act differently if there exists vagueness or ambiguity in the probabilities of the prospect. For more on source dependence, see Heath and Tversky [29]. Expected utility theory does not account for source dependencies. Prospect theory also does not take this into account; however, cumulative prospect theory will.

Many studies have confirmed the above risk attitudes, see for example, [18], [17], [24], [35], [16], and [26]. Expected utility theory claims to be a model of rational choice but it can be argued that much of the above behaviour is quite rational. Clearly, a new model is needed to incorporate loss aversion, risk seeking, nonlinear preferences and source dependence.

1.3 Prospect theory

Prospect theory tries to take into account the above behavioural anomalies that expected utility theory is unable to incorporate. The main changes to expected utility theory are summarized below:

- Outcomes are defined in terms of gains and losses relative a reference point
- There is a nonlinear probability decision weighting function that distorts individual probabilities

Kahneman and Tversky do not specifically include source dependence into prospect theory, however cumulative prospect theory does take this into account, which will be looked at in the next chapter. This theory looks at risky prospects where the probabilities are known by the decision maker explicitly. According to prospect theory the value of a prospect, $V(\mathbf{x}, \mathbf{p})$ is given by the following following formula,

$$V(\mathbf{x}, \mathbf{p}) = \pi(p_1)v(x_1) + \pi(p_2)v(x_2) + \ldots + \pi(p_n)v(x_n) \quad (1.3.1)$$

$$= \sum_{i=1}^{n} \pi(p_i)v(x_i) \quad (1.3.2)$$

All outcomes x_i are given in terms of a reference point which for simplicity is usually given the value $x_0 = 0$, and all other outcomes are defined as gains or losses in terms of the reference point. Positive values of x_i are considered gains, and negative values are losses. For example, a typical reference point is an individuals current asset position. A decision weighting function is introduced and is denoted by $\pi(p_i)$; this gives the probability distortion of every individual probability by either overweighting p_i or underweighting it. However, $\pi(p_i)$ is not a probability measure and $\sum \pi(p_i)$ typically does not add to unity. The decision weights have the following property: $\pi(0) = 0$, $\pi(1) = 1$. This makes sense intuitively, because individuals would not distort impossibility and certainty.

A value function is also introduced, this is similar to the utility function in expected utility theory. It is denoted by v, and it assigns to every outcome x_i a number $v(x_i)$ in terms of the reference point. It is assumed that $v(0) = 0$, which is intuitively clear because the value of the reference point should be neutral. As well, $v(x)$ is assumed to be a continuous, strictly increasing function [17]. Similar to expected utility theory, under prospect theory a prospect $(\mathbf{x}, \mathbf{p})$ is preferred to another prospect, $(\mathbf{y}, \mathbf{q})$, if $V(\mathbf{x}, \mathbf{p}) > V(\mathbf{y}, \mathbf{q})$ and is indifferent if $V(\mathbf{x}, \mathbf{p}) = V(\mathbf{y}, \mathbf{q})$. Preferences are determined jointly by a utility function that evaluates the subjective value of an outcome in terms of a reference point, and by the decision weights that capture an individual's attitude towards risk. [32]

Consider now the shape of the value function, $v(x)$. We defined the value function in terms of a reference point and the outcomes were either considered gains or losses. The reference point usually corresponds to the current asset position, where gains and losses are the actual increase or decrease in wealth. However, the reference point need not be an individuals current wealth position, the reference point can be affected by the composition of the prospects being offered, and by the expectations of the individual making the decision [17]. We define the value function to exhibit diminishing marginal value as the monetary value of the outcomes move farther away from the reference point. This is referred to as diminishing sensitivity. In the gains domain this implies concavity, $v''(x) < 0$. This means that a difference in salary from \$20,000 to \$30,000 has a much bigger impact than \$70,000 to \$80,000. In the losses domain this implies convexity, $v''(x) > 0$. The difference of a loss of \$100 to \$200 is much more significant than the difference between \$1,000 to \$1,100. In other words, the impact of every dollar in debt gets smaller as your debt increases.

The property of loss aversion is also incorporated into the shape of the value function. Since losses loom for longer than gains we characterize the value function as steeper for losses than for gains. This means that $v'(x)$ for $x \geq 0$ is less than $v'(x)$ for $x \leq 0$. Figure 1.1 gives a graphical example of the above properties; an asymmetrical s-shaped function. In summary these properties are reference point dependence, diminishing sensitivity, and loss aversion. There are, however, cases where the value function has characteristics not described by these properties. For example there may be a steep increase in an individuals value function as x approaches $30,000 if that individuals needs $30,000 to enroll in university [17].

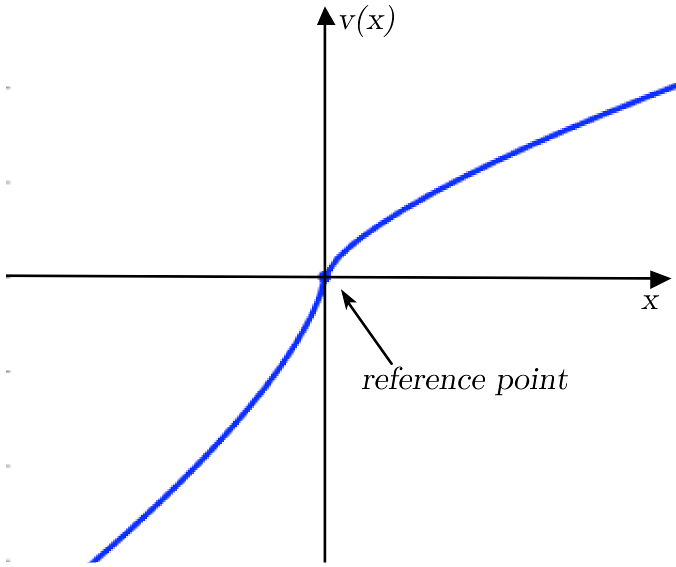

Figure 1.1: The value function, which is defined in terms of reference point

Prospect theory proposes to distort individual probabilities using a decision weighting function; it is a monotonic transformation of outcome probabilities [31]. This seems like the intuitive and simplest way of extended expected utility theory to include nonlinear preferences. Prospect theory, by including distortions of probabilties, reference dependence, diminishing sensitivity, and loss aversion can explain major deviations from expected utility theory, but there are still theoretical problems with this theory. The functional form does not always satisfy stochastic dominance and the theory cannot be easily extended to prospects with a large number of outcomes. Stochastic dominance is used as a way of comparing prospects. A given prospect $(\mathbf{x}, \mathbf{p})$ has first-order stochastic dominance over $(\mathbf{y}, \mathbf{q})$ if for every outcome the first prospect gives a higher probability of receiving an outcome greater than or equal to that outcome. If we look at this in terms of the cumulative probability distributions of both prospects then $(\mathbf{x}, \mathbf{p})$ with cumulative probability distribution F_x dominates $(\mathbf{y}, \mathbf{p})$ with cumulative probability distribution F_y if for every outcome $F_x \leq F_y$ [4]. Stochastic dominance implies that a shift in probability mass from worse outcomes to better outcomes would lead to a better prospect, in other words an increased value of V [11].

Now, to see how prospect theory is not easily extended to a large number of outcomes consider the following prospect: $(-10, 0.05; 0, 0.05; 10, 0.05; \ldots; 180, 0.05)$. As was stated earlier, small probabilities tend to be overweighted, and this therefore implies that $\pi(0.05) > 0.05$. This means that, assuming a concave value function, $v(x)$, for positive values of x,

$$V(\mathbf{x}, \mathbf{p}) \geq v(E[\mathbf{x}]) = v(85). \tag{1.3.3}$$

where \$85 is the expected value of the propect. This is not realistic, since most individuals would prefer the \$85 for certain. This comes as the result of overweighting all outcomes, which is why prospect theory cannot be extended to prospects with a large number of outcomes [11]. Cumulative prospect theory, to eliminate these problems, transforms cumulative probabilities instead of individual probabilities and will be discussed in greater detail in the next chapter.

Chapter 2

Cumulative Prospect Theory

2.1 Introduction to Cumulative Prospect Theory

In 1992, Tversky and Kahneman [31] presented a new version of prospect theory called cumulative prospect theory. Similar to prospect theory, this new theory incorporates a value function and a probability distortion function. The value function, like prospect theory, exhibits reference dependence, diminishing sensitivity, and loss aversion. However, cumulative prospect theory modifies prospect theory by applying the probability distortions to the cumulative probabilities as opposed to the individual probabilities, in order to include nonlinear preferences. This formulation is called rank-dependence and was first proposed by Quiggin [23], The advantage being that it satisfies stochastic dominance, see [4] for an analysis of this property and cumulative prospect theory, and has the ability to be applied to prospects with any number of outcomes. Cumulative prospect theory also allows for different distortion functions while in the domain of gains or in the domain of losses.

Cumulative prospect theory generalizes prospect theory to decisions not only made under risk but to those made under uncertainty. For the purposes of this paper we distinguish between a decision made under risk to one that is made under uncertainty. A choice made under risk refers to a choice where the decision maker knows or is given explicit probabilities for each outcome of the prospect. A choice under uncertainty is when the decision maker does not know the probability of each outcome or the probabilities are vague in some way or there is some kind of ambiguity. This allows the model to take into account source dependence. In summary, cumulative prospect theory includes loss aversion, risk seeking, nonlinear preferences, and source dependence but satisfies stochastic dominance, and can easily be extended to prospects with a large number of outcomes.

The next sections describe the mathematical representation of the theory proposed by Tversky and Kahneman [31]. We will first look at the model for decisions under risk then describe the model under uncertainty.

2.1.1 Decisions under risk

When making a decision under risk the probabilities of each outcome of a prospect is known to the decision making before the decision is made. Similar to prospect theory we denote a prospect by $(\mathbf{x}, \mathbf{p})$, where $\mathbf{x}$ are the outcomes of the prospect and $\mathbf{p}$ are their corresponding probabilities. We define the reference point as $x_0 = 0$ and all other outcomes are defined in terms of this reference point. However, we now give negative subscripts to negative outcomes or losses, and positive subscripts to positive outcomes or gains. A prospect with $n + m + 1$ outcomes is therefore given by $(x_{-m}, p_{-m}; \ldots; x_n, p_n)$, where $n, m \geq 0$, and $x_{-m} \leq \ldots \leq x_n$. The ordering of outcomes is important for the rank-dependent formulation of cumulative prospect theory. For convenience we denote a prospect by f and f^+ is the positive part of the prospect, $(x_1, p_1; \ldots; x_n, p_n)$, and f^- is the nonpositive part, $(x_{-m}, p_{-m}; \ldots; x_0, p_0)$. The value of a prospect is given by the following equation,

$$V(\mathbf{x}, \mathbf{p}) = V(f) = V^-(f^-) + V^+(f^+), \tag{2.1.1}$$

which is separated into two parts in terms of gains and losses. If all the outcomes in a prospect are all positive then $V(f) = V^+(f^+)$ and if the outcomes are all negative then $V(f) = V^-(f^-)$. The following equation defines the value for the positive outcomes,

$$V^+(f^+) = \sum_{i=1}^{n} \pi_i^+ v(x_i). \tag{2.1.2}$$

and similarly the following equation defines the value of the negative outcomes.

$$V^-(f^-) = \sum_{i=-m}^{0} \pi_i^- v(x_i), \tag{2.1.3}$$

where, like prospect theory it assumed that there exists a strictly increasing value function $v : x \to \mathbb{R}$ satisfying $v(x_0) = v(0) = 0$,

$$\pi^+(f^+) = (\pi_1^+, \ldots, \pi_n^+) \tag{2.1.4}$$

$$\pi^-(f^-) = (\pi_{-m}^-, \ldots, \pi_0^-), \text{ and} \tag{2.1.5}$$

decision weights or probability distortions for gains are given by the following formulas

$$\pi_n^+ = w^+(p_n) \tag{2.1.6}$$

$$\pi_i^+ = w^+(p_i + \ldots + p_n) - w^+(p_{i+1} + \ldots + p_n), \text{ for } 0 \leq i \leq n-1 \tag{2.1.7}$$

$$= w^+\left(\sum_{j=i}^{n} p_j\right) - w^+\left(\sum_{j=i+1}^{n} p_j\right), \tag{2.1.8}$$

and similarly the decision weights for the losses are given by

$$\pi^-_{-m} = w^-(p_{-m}) \tag{2.1.9}$$

$$\pi^-_i = w^-(p_{-m} + \ldots + p_i) - w^-(p_{-m} + \ldots + p_{i-1}), \text{ for } 1 - m \le i \le 0 \tag{2.1.10}$$

$$= w^-\left(\sum_{j=-m}^{i} p_j\right) - w^-\left(\sum_{j=-m}^{i-1} p_j\right). \tag{2.1.11}$$

where w^+ and w^- are strictly increasing functions satisfying $w^+(0) = w^-(0) = 0$ and $w^+(1) = w^-(1) = 1$, since if something is impossible it should not impact an individual's preference and if something certainly will happen then the effect should be the value that outcome is given. We will refer to π_i as the decision weights or the probability distortion functions, and w^+ and w^- as the decision weighting functions.

Notice that the decision weight, π^+_i, of a gain is the difference between the decision weighting function for gains, denoted by w^+, for the cumulative probabilities of outcomes that are at least as good as x_i and the decision weighting function for the cumulative probabilities of outcomes that are strictly better than x_i. The decision weight, π^-_i, of a loss is the difference between the decision weighting function for losses, denoted by w^-, for the cumulative probabilities of outcomes that are at least as bad as x_i and the decision weighting function for the cumulative probabilities of outcomes that are strictly worse than x_i. The decision weights can thus be interpreted as the marginal contribution of each respective outcome [31]. If w is additive, then π_i is simply the probability of each outcomex_i. For both positive and negative prospects the decision weights add to one, but for mixed prospects the sum can either be smaller or greater than 1, due to the fact that the decision weights may be defined by different functions. One should also be aware of the fact that decision weights are not probabilities; they should not be understood as measures of belief [17]. Finally, we have the value of a prospect given by,

$$V(\mathbf{x}, \mathbf{p}) = w(p_{-m})v(x_{-m}) + \sum_{i=-m+1}^{0} \left[w^-\left(\sum_{j=-m}^{i} p_j\right) - w^-\left(\sum_{j=-m}^{i-1} p_j\right)\right] v(x_i) \tag{2.1.12}$$

$$+ w(p_n)v(x_n) + \sum_{i=1}^{n-1} \left[w^+\left(\sum_{j=i}^{n} p_j\right) - w^+\left(\sum_{j=i+1}^{n} p_j\right)\right] v(x_i). \tag{2.1.13}$$

Again, similar to expected utility theory we assign each prospect a number $V(\mathbf{x}, \mathbf{p})$ such that $(\mathbf{x}, \mathbf{p})$ is preferred to $(\mathbf{y}, \mathbf{q})$ if $V(\mathbf{x}, \mathbf{p}) > V(\mathbf{y}, \mathbf{q})$ and is indifferent if $V(\mathbf{x}, \mathbf{p}) = V(\mathbf{y}, \mathbf{q})$. In the next subsection, decisions under uncertainty will be addressed.

2.1.2 Decisions under uncertainty

Decision under uncertainty relates to the case where the decision maker is unaware, or the probability associated with an outcome(s) in the prospect is vague. This allows one to take into account source dependencies. We assume that each the outcome, x_i, is the result of the uncertain event A_i taking place, where the event A_i is a subset of the state space denoted by S. In other words, S is the set of all possible states in the universe [28]. An uncertain prospect is denoted by $(x_{-m}, A_{-m}; \ldots; x_n, A_n)$ and for short $(\mathbf{x}, \mathbf{A})$. For decisions under risk, we have as before

$$V(f) = V^+(f^+) + V^-(f^-) \tag{2.1.14}$$

$$= \sum_{i=-m}^{0} \pi_i^- v(x_i) + \sum_{i=1}^{n} \pi_i^+ v(x_i), \tag{2.1.15}$$

but because probabilities are unknown we are not able to use a transformation of the probabilities- we must define the weighting function for gains W^+ and for losses W^- as a mapping that assigns to each event A_i in S a number between 0 and 1 satisfying $W^-(\emptyset) = W^-(\emptyset) = 0$, $W^+(S) = W^-(S) = 1$, and $W^+(A_i) \leq W^+(A_j)$ and $W^-(A_i) \leq W^-(A_j)$ whenever $A_i \supset A_j$. They are defined as,

$$\pi_n^+ = W^+(A_n) \tag{2.1.16}$$

$$\pi_i^+ = W^+(A_i \cup A_{i+1} \cup \ldots \cup A_n) - W^+(A_{i+1} \cup \ldots \cup A_n), \text{ for } 0 \leq i \leq (n-1) \tag{2.1.17}$$

and

$$\pi_{-m}^- = W^-(A_{-m}) \tag{2.1.18}$$

$$\pi_i^- = W^-(A_{-m} \cup \ldots \cup A_{i+1} \cup A_i) - W^-(A_{-m} \cup \ldots \cup A_{i-1}) \text{ for } (1-m) \leq i \leq 0. \tag{2.1.19}$$

The decision weighting functions W^+ and W^- are defined by the above properties, and can be derived from preferences, but they are not directly observable [34]. Therefore, we will assume a two-stage decision process, proposed by Tversky and Fox [28], for decisions under uncertainty. The two-stage process is defined as follows: first the decision maker makes a judgment about the perceived probability, $P(A_i)$ of an event, A_i, then transforms this value by the risky decision weighting function $w^-(P(A_i))$ or $w^+(P(A_i))$. Tversky and Fox [28], Fox *et al* [13], and Kilka and Weber [19] found evidence of this two-stage decision making process.

The axiomatisation of cumulative prospect theory for decisions under risk and for decisions under uncertainty are presented by Chateauneuf and Wakker [8], and Wakker and Tversky [34], respectively. The next section will look at the parametric forms of the cumulative prospect theory.

2.2 Parametric Forms

Under cumulative prospect theory preference depends jointly on the value function, $v(x)$, and the weighting function, $w(p)$. The parametric forms of the value function need to display the fourfold risk pattern: risk seeking for gains and risk aversion for losses of low probability combined with risk aversion for gains and risk seeking for losses of high probability. The introduction of decision weights makes the model become quite complicated and it is sometimes unclear what effects the value function and the decision weighting function will have together. The superscripts of + and − will be omitted when it is clear that we are talking about just gains, losses, or the analysis is the same for both gains and losses. Next, the parametric form of the value function will be described, then the form of the decision weighting function, and lastly the effects of the value and the weighting function will be discussed.

2.2.1 The value function

The value function, $v(x)$, for cumulative prospect theory exhibits the same properties as for prospect theory: reference dependence, diminishing sensitivity, and loss aversion. This means that $v(x)$ is concave above the reference point, so $v'' \leq 0, x \geq 0$, and convex below the reference point, $v'' \geq 0, x \leq 0$. This reflects diminishing sensitivity, which means that the impact of changes in the domain of gains and losses diminishes when the distance from the reference point increases. The value function is also steeper for losses than for gains, i.e. $v'(x) < v'(-x)$ for $x \geq 0$, since losses loom for longer than gains. The following equation is the parametric form of the value function suggested by Kahneman and Tversky [31].

$$v(x) = \begin{cases} x^{\alpha}, & \text{if} \quad x \geq 0 \\ -\lambda(-x)^{\beta}, & \text{if} \quad x < 0 \end{cases} \tag{2.2.1}$$

where $\lambda \geq 1$ $0 \leq \beta \leq 1$, and $0 \leq \alpha \leq 1$. As α and β increase the effect of diminishing sensitivity decreases, and as λ increases the degree of loss aversion increases.

2.2.2 The Decision Weights

The decision weighting function takes cumulative probabilities and weights them nonlinearly. The parametric form of the decision weighting function can, therefore, incorporate nonlinear preferences and the fourfold risk pattern.

The principle of diminishing sensitivity applies not only to the value function but also to the weighting functions. The difference is that the sensitivity to changes in probability diminish as probability moves away from the boundary points of impossibility and certainty. This means that a change in the probability of an outcome from 0.1 to 0 and 0.9 to 1 has more of impact then a change from 0.5 to 0.6. Also, small probabilities tend to be overweighted and moderate to high probabilities tend to underweighted These properties gives rise to a function which is concave near 0 and convex near 1 (an inverse s-shape). Moreover, we can say this implies bounded subadditivity. A function $w(p)$ exhibits bounded subadditivity if there exists constants ϵ_1 and ϵ_2 such that

$$w(q) \geq w(p+q) - w(p) \text{ whenever } p+q \leq 1-\epsilon_1 \tag{2.2.2}$$

and

$$1 - w(1-q) \geq w(p+q) - w(p) \text{ whenever } p \geq \epsilon_2 \tag{2.2.3}$$

The following equations give the parametric form proposed by Kahneman and Tversky [31].

$$w^+(p) = \frac{p^\gamma}{(p^\gamma + (1-p)^\gamma)^{\frac{1}{\gamma}}}, \tag{2.2.4}$$

and

$$w^-(p) = \frac{p^\delta}{(p^\delta + (1-p)^\delta)^{\frac{1}{\delta}}}. \tag{2.2.5}$$

The decision weighting functions for gains and losses have the general shape, the inverse s-shape, but, the weighting function for losses tends to higher and less curved [31]. The parametric forms given by equation 2.2.4 and equation 2.2.5 are given by plotted in Figure 2.1 for different values of γ. As γ increases, the degree of curvature increases.

Now, consider the prospect referred to in the previous chapter, $(-10, 0.05; \ldots; 180, 0.05)$. Under cumulative prospect theory only extreme outcomes are overweighted. In this example, 180 has a decision weight given by $w^+(0.05)$, and the other extreme outcome -10 has a decision weight given by $w^-(0.05)$. These outcomes are overweighted in accordance with the shape of $w(p)$. The moderate outcomes receive smaller decision weights, for example the 100 outcome receives a decsion weight of $w^+(0.45) - w^+(0.4)$ and this outcome is underweighted according to the invserve s-shape. This agrees with the principle of diminshing

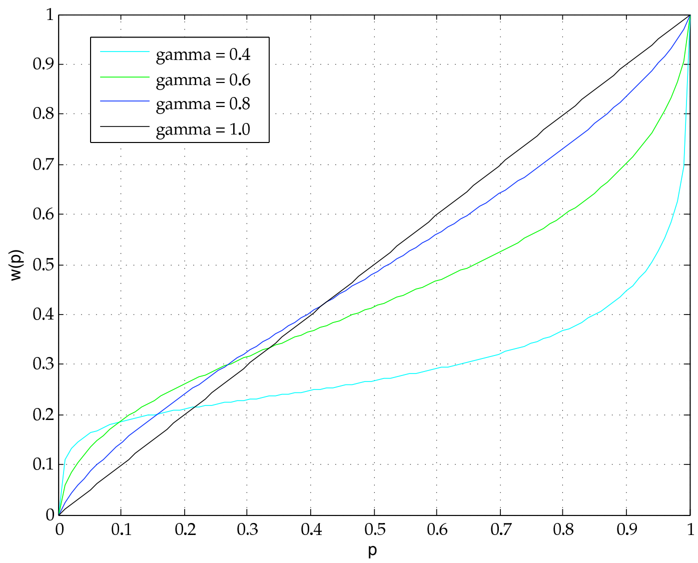

Figure 2.1: The plot of the decision weighting function $w(p)$

sensitivity with respect to the outcomes. This means that individuals are more sensitive to extreme outcomes and less sensitive to moderate outcomes; experimental evidence also shows this is the case [11].

The value function and the decision weighting function incorporate some important properties of behaviour, but it is not entirely obvious what the properties yield when these two functions are combined together in the cumulative prospect theory framework. If the value function is linear then an individual's preferences are solely determined by the probability weighting function. In this case the inverse s-shape of the decision weights give rise to the fourfold risk pattern. The overweighting of outcomes with small probability implies risk seeking in the domain of gains and risk aversion in the domain of losses. The underweighting of moderate to large probabilities implies risk aversion in the domain of gains and risk seeking in the domain of losses. However, the value function exhibits diminishing sensitivity, therefore the probability weighting function must compensate for the shape of the value function, which sometimes works in favour and sometimes against the fourfold pattern [22].

Cumulative prospect theory not only corrects the theoretical problems with prospect theory by incorporating rank-dependence but incorporates source dependence, and separate decision weights for gains and losses. Expected utility theory is also a special case of cumulative prospect theory, therefore, one does not have to fully abandon expected utility when adopting cumulative prospect theory. The next section will describe other parametric forms that have been proposed.

2.3 Other Parametric Forms

The decision weighting function proposed by Kahneman and Tversky [31] is not the only parametric form that exhibits properties consistent with the principles of cumulative prospect theory more specifically, the inverse s-shape. The following decision weight was proposed by Gonzalez and Wu [15] ,

$$w(p) = \frac{\eta p^{\gamma}}{\eta p^{\gamma} + (1-p)^{\gamma}}. \tag{2.3.1}$$

The function incorporates two parameters: $\eta, \gamma > 0$. It was derived by separating out a parameter for the curvature and a parameter for the elevation of the inverse s-shape. The curvature parameter γ represents the degree of diminishing sensitivity: the degree of curvature decreases as γ increases. The elevation parameter η represents the attractiveness of the bet: the elevation increases as η increases.

Another form of the probability weighting function was proposed by Prelec [22].Other parametric forms were chosen based on empirical evidence, but Prelec used the empircal evidence to derive the parametric form. The following equation is the one-parameter form,

$$w(p) = e^{-(-logp)^{\gamma}}, \tag{2.3.2}$$

where $\gamma > 0$ and the point at which the weighting function crosses the line $w(p) = p$ is fixed at $\frac{1}{e} \approx 0.3679$. The inflection point for the parametric equations typically occur at $p < 0.4$The function exhibits the inverse s-shape, concave for lower probabilities and convex for the upper probabilities. Similarly for γ in Kahneman and Tversky's model, γ controls how exaggerated the overweighting and underweighting are.

This paper, however, will focus on the weighting function proposed by Kahneman and Tversky [31] described in the previous section. See [22] for a derivation and detailed analysis of the properties of the above forms.

There are also many alternatives that one could use for the value function, for example a logarithmic fuction, however most literature uses a power function. See [27] for an investigation of which functional forms in which combinations have the most explanatory power. The next section will overview methods of constructing the value function and weighting function using a parameter-free approach.

2.4 Parameter-free Elicitation of the Functional

The assumption of a defined parameteric form has the advantage that estimated value functions and decision weights are smooth functions and have a nice form, but parameter estimates may not actually fit the data. It is essentially forcing the data to show what we want it to show by choosing the parameters to fit the data. The advantage of parameter-free methods is that they do not assume a parametric form and do not impose any restrictions *a priori*, because fewer assumptions are made. The disadvantage of these methods is that estimates tend to be noisier. Parameter-free elicitation of the cumulative prospect theory functional form has been investigated by Gonzalez and Wu [15], Adbellaoui [2], and Bleichrodt and Pinto [5].

Gonzalez and Wu [15] present a method of deducing the value function and the probability weighting function by using tradeoffs between gambles and certainty equivalents. A certainty equivalent, C, refers to the value for certain that an individual is indifferent between a prospect and that value for certain, i.e.

$$V(\mathbf{x}, \mathbf{p}) = v(C). \tag{2.4.1}$$

Gonzalez and Wu found evidence that the value function, $v(x)$, is concave for gains, and the probability weighting function, $w(p)$, exhibits the inverse s-shape. However, they also found that there is a considerable amount of variability among individual functional forms, which is natural to assume that findings at the aggregate level may not hold for all individuals.

Adbellaoui [2] proposed a two-step method that elicits the value function and decision weighting function using the tradeoff method first described by Wakker and Deneffe [33]. This method looks at the indifference between two gambles rather than certainty equivalents to determine the functional forms. Adbellaoui's findings were consistent with previous assertions that the value function is concave for gains and convex for losses, and the probability weighting functions display lower and upper subadditivity. Bleichrodt and Pinto [5] also used the trade-off method and found that there is significant evidence that individuals exhibit nonlinear probability weighting functions. This was seen both at the individual level and the aggregate level. Again, there was also evidence that the decision weighting function displayed lower and upper subadditivity consistent with the inverse s-shape.

These non-parametric estimations coupled with the parametric estimations are a good way to validate the use of the parametric forms. They can be used a check to make sure the parametric forms are the best way to capture the data, but the functional forms provide a convenient and easy way of predicting out of sample situations. These studies all found evidence that is consistent with Kahneman and Tversky's functional form, therefore, we can be more confident in the parametric assumption. The next chapter will look at the parameter estimation using data from the lottery.

Chapter 3

Parameter Estimation

Estimates of the parameters for Kahneman and Teversky's functional forms have been approximated by Kahneman and Tversky [31], Cramerer and Ho [7], and Gonzalez and Wu [35]. These estimates were calculated using data collected from hypothetical experiments where test subjects were asked questions about their preferences for prospects involving one or two outcomes at most. These textbook gambles do not accurately resemble real choices that individuals are face [36]. This is why the objective here is to use data from a real world context to see whether or not cumulative prospect theory captures the behaviour and whether or not the estimated parameters can be used to predict behaviour in out of sample situations. We decided to look at data from the lottery market. Situations involving very extreme outcomes is rarely addressed in literature, as well as prospects with more than two or three outcomes. The advantage of looking at extreme outcomes is that we can easily find the degree of overweighting of small probabilities.

3.1 The Lottery Market

We use lottery data retrieved from a lottery called Mark 6 that is based in Hong Kong to estimate the parameters in cumulative prospect theory using Kahneman and Tversky's parametric value function (equation (2.2.1)) and weighting function (equation (2.2.4)). We assume a single preference structure for all individuals. The purpose of this is so that we can obtain a model that will represent the decisions made by an average person makes decisions under risk and uncertainty. Obviously, individuals do not all behave in the same way but for analytical tractability we assume that there exists a single representative agent that reflect the market as a whole [7]. Also, because we are dealing with data from the lottery market we do not have the luxury of knowing any information about individual preferences we only have the aggregate data for example, total sales, the value of prizes, and the number of price winners.

The cost of a ticket is 5 Hong Kong dollars (HKD). All monetary currency for this section will be quoted in HKD (Hong Kong dollar is pegged at the U.S. dollar to $1USD = 7.8\ HKD$). Each ticket allows the bettor one chance to pick six numbers from 1 to 49. For each draw six numbers are chosen and in addition an extra number is also drawn. There are seven divisions of prizes one could win. To win the first division, all six numbers must be picked, in the second divison five numbers plus the extra number, third division any five numbers excluding the extra number, the fourth prize is given to those who have any four numbers plus the extra number, in the filth division, any four of the numbers excluding the extra number, the sixth prize is any three numbers plus the extra number, and finally the last prize division is given to those who get any three numbers exlcuding the extra number. All information on Mark 6 is detailed on their official website, see [1]. The following calculates the theoretical probability of winning each division prize.

$$P(\text{Winning } 1^{\text{st}} \text{ division prize}) = p_1 = \frac{\binom{6}{6}}{\binom{49}{6}} = \frac{1}{13,983,816} \approx 7.1511 \times 10^{-8}$$

$$P(\text{Winning } 2^{\text{nd}} \text{ division prize}) = p_2 = \frac{\binom{6}{5}\binom{1}{1}}{\binom{49}{6}} = \frac{1}{2,330,636} \approx 4.2907 \times 10^{-7}$$

$$P(\text{Winning } 3^{\text{rd}} \text{ division prize}) = p_3 = \frac{\binom{6}{5}\binom{42}{1}}{\binom{49}{6}} = \frac{3}{166,473} \approx 9.0104 \times 10^{-6}$$

$$P(\text{Winning } 4^{\text{th}} \text{ division prize}) = p_4 = \frac{\binom{6}{4}\binom{1}{1}\binom{42}{1}}{\binom{49}{6}} = \frac{15}{332948} \approx 4.5052 \times 10^{-5}$$

$$P(\text{Winning } 5^{\text{th}} \text{ division prize}) = p_5 = \frac{\binom{6}{4}\binom{42}{2}}{\binom{49}{6}} = \frac{12,915}{13,983,816} \approx 9.2357 \times 10^{-4}$$

$$P(\text{Winning } 6^{\text{th}} \text{ division prize}) = p_6 = \frac{\binom{6}{3}\binom{1}{1}\binom{42}{2}}{\binom{49}{6}} = \frac{4,305}{3,495,954} \approx 1.2314 \times 10^{-3}$$

$$P(\text{Winning } 7^{\text{th}} \text{ division prize}) = p_7 = \frac{\binom{6}{3}\binom{42}{3}}{\binom{49}{6}} = \frac{4,100}{249,711} \approx 1.6419 \times 10^{-2}$$

The first to third divison prize values vary depending on the sales of the tickets; however the first division prize is guaranteed to be allocated at least $\$5,000,000$. If no one wins the first or second division prize then it gets carried over to the next draw's first division prize. If no one wins the third division prize then it gets spread into the first and second division prizes of that same draw. If more than one person wins the first, second, or third divsion then the prize is divided equally among the holders of the winning ticket. The last four division prizes are fixed at $4800, $320, $160, and $20.

The market price of the lottery ticket is $5. The first, second, and third division prizes vary from draw to draw, according to the sales and rollover from past prizes. This can increase the value of the first division prizes dramatically. We want to split the data up into classes based on the amount of rollover there is from previous unwon prizes. This means that we will define each class to be given by the number of prior draws that had no winners. This varies from 0 to 6. The past winners of lottery tickets are usually publicised and individuals buying lottery tickets will be aware of which class the ticket enters into. [9] These classes will give us the data needed to find the parameters that reflect the market price of the lottery prospect. Table 3.1 gives the averaged data for each class.

Number of draws prior with no win	Number of samples	1st Division average prize	2nd Division average prize	3rd Division average prize
0	100	$12,868,476.80	$594,754.65	$34,371.85
1	69	$14,526,918.33	$731,742.61	$35,359.35
2	45	$13,203,704.00	$462,040.22	$33,144.89
3	33	$14,136,913.94	$643,626.97	$34,419.09
4	18	$12,794,725.56	$499,868.89	$32,461.11
5	7	$24,845,828.57	$686,297.86	$41,650.71
6	7	$17,485,475.00	$948,558.57	$43,402.86
> 6	0	$0.00	$0.00	$0.00
All Samples	279	$13,894,315.66	$618,060.35	$34,709.70

Table 3.1: Average data for each class. Note that the averages in the bottom row were calculated using a weighted average with respect to the number of samples.

We will assume that individuals purchase only one lottery ticket. Obviously, some people buy more than one lottery ticket, however since we do not have the data to take this into account we assume that every ticket is purchased by a different person.

3.1.1 Reference Point of the Lottery Prospect

The are two main choices for the reference point when looking at the lottery prospect. They are as follows:

1. The price of the lottery ticket, or
2. The current asset position of the ticket purchaser

We will first look at the case of using the price of the lottery ticket as the reference point. Consider the expected value of the prospect calculated using the weighted averages from Table 3.1. Note that for convenience we denote the first division prize by x_1 and the second division prize by x_2, and so on, instead of assigning the largest outcome the largest subscript. The expected value of the lottery prospect is

$$E[\mathbf{x}] = p_1 x_1 + \ldots + p_7 x_7 \tag{3.1.1}$$

$$\approx 2.92 \tag{3.1.2}$$

Since expected utility theory assumes risk aversion or equivalently a concave utility function, the above equation implies, using Jensen's inequality, that

$$E[u(\mathbf{x})] \leq u(E[\mathbf{x}]) \approx u(2.92) \leq u(5) \tag{3.1.3}$$

This means that the utility of the price of the lottery ticket is greater than the expected utility of the lottery prospect, thus according to expected utility theory individuals would not buy lottery tickets, which obviously is not the case in the real world. Under the cumulative prospect theory framework individuals would be willing to buy a lottery ticket if

$$V(\mathbf{x}, \mathbf{p}) \geq v(5) \tag{3.1.4}$$

$$\Rightarrow V(\mathbf{x}, \mathbf{p}) - v(5) \geq 0 \tag{3.1.5}$$

where

$$V(\mathbf{x}, \mathbf{p}) = w(p_1)v(x_1) + [w(p_1 + p_2) - w(p_1)]v(x_2) \tag{3.1.6}$$

$$+ \ldots + [w(p_1 + \ldots + p_7) - w(p_1 + \ldots + p_6)]v(x_7), \tag{3.1.7}$$

and $v(x)$ is given by equation (2.2.1). In this approach we are treating the price of the ticket as the certainty equivalent of the prospect. This certainty equivalent was determined by the market. This means that an individual or the single representative economic agent is indifferent between \$5 and the lottery prospect. This approach has the advantage that we are only in the domain of gains and therefore only have two parameters that need to be estimated, α and γ. To find the parameters that reflect the certainty equivalent found in the market we set $V(\mathbf{x}, \mathbf{p}) = v(5)$ and find α and γ that best fits this equation.

If we instead take as the reference point the initial assets of the purchaser of the lottery ticket then we look at this prospect as a large chance of losing a small amount plus a very small chance of winning a large amount. Clearly expected utility theory cannot account for the purchase of a lottery ticket since the expected value for this prospect is negative. Under cumulative prospect theory an individual would be wiling to take on the lottery prospect if

$$V(\mathbf{x}, \mathbf{p}) \geq 0 \tag{3.1.8}$$

where,

$$V(\mathbf{x}, \mathbf{p}) = w^{-}(1 - [p_1 + \ldots + p_7])v(-5) \tag{3.1.9}$$

$$+ w^{+}(p_1)v(x_1 - 5) + [w^{+}(p_1 + p_2) - w^{+}(p1)]v(x_2 - 5) \tag{3.1.10}$$

$$+ \ldots + [w^{+}(p_1 + \ldots + p_7) - w^{+}(p_1 + \ldots + p_6)]v(x_7 - 5) \tag{3.1.11}$$

This prospect is both in the domain of gain and loss, therefore, there are five parameters, α, γ, δ, λ, and β, that need to be estimated for this approach (see equations (2.2.1), (2.2.4), (2.2.5)). We will, thus, focus on the first approach, because we do not have enough data to accurately estimate five parameters.

Theoretically, one can see how cumulative prospect theory can explain the purchase of a lottery ticket, even though it is an unfair prospect, by the over weighting of small probabilities. However, the value function, $v(x)$, experiences diminishing value as x increases, therefore, the overweighting has to be pronounced enough to compensate for this. Our estimations will show us whether or not the parametric functional form propsed by Kahneman and Tversky holds up in this kind of extreme event. The next section will describe the methods of parameter estimation.

3.2 Method of Estimation

The estimates of the parameters α and γ were calculated first using local search optimization, and then a nonlinear least squares approach was used to gain a more refined solution. However, when the parameters needed to be constrained, only the local search optimization was used. The following details first the local search optimization then the nonlinear least squares approach. Both approaches were implemented using Matlab, and both codes can be found in the appendix.

3.2.1 Local search optimization

In order to ensure the fourfold risk pattern the search space for both parameters α and γ is restricted to be between 0 and 1. Since the search space is quite small, it is easy to find the parameters that give the smallest mean squared error by using a local search optimization. This is accomplished by creating a vector of possible values of α and γ and testing out each combination of α and γ. We use the weighted least squares because the frequency of the data varies and we want to give more importance to those samples or classes that were observed more frequently in the lottery market. Refer back to Table 3.1 to see the frequencies. We then refine the search around this parameter combination, and keep refining until a sufficient error tolerance is reached.

This method has the advantage of being simple to implement and it also allows one to constrain the values of the parameters. However, this method has the disadvantage of being very inefficient and it is unclear whether you have found global minimum or a local minimum. The minimum value could be dependent on the initial state space. This is why we have used a nonlinear least squares approach as well.

3.2.2 Nonlinear Least Squares

The local search optimization is the simplest way to approximate the parameters; however, nonlinear least squares is more efficient, gives a more refined solution, and decreases the error in the local search optimization. The nonlinear least squares method used is called Newton's method which is used to estimate the parameters in the model. The method follows that presented in by Scales [25] and Aster *et al* [3]. This method minimizes the weighted sum of the square of the residual.

Let m be the number of samples, l be the number of parameters in the model, and $\boldsymbol{\theta}$ be the vector of parameters to be estimated. The residual $r_i(\boldsymbol{\theta})$ is given by

$$r_i(\boldsymbol{\theta}) = V(\mathbf{x}^{(i)}, \mathbf{p}; \boldsymbol{\theta}) - v(C; \boldsymbol{\theta}), \quad i = 1, \dots m \tag{3.2.1}$$

where $V(\mathbf{x}^{(i)}, \mathbf{p}; \boldsymbol{\theta})$ is the value of the prospect evaluated for every sample i, $v(C; \boldsymbol{\theta})$ is the value of the certainty equivalent for the prospect, and $\mathbf{x}^{(i)} = (x_1^{(i)}, x_2^{(i)}, \dots)^T$ is the value of each outcome for each sample. Now, define $R(\boldsymbol{\theta})$ to be the weighted sum of the square of the residuals,

$$R(\boldsymbol{\theta}) = \sum_{i=1}^{m} b_i r_i(\boldsymbol{\theta})^2, \tag{3.2.2}$$

where b_i is the weight associated with each sample i. If we let $f(\boldsymbol{\theta}) = \sqrt{b} r(\boldsymbol{\theta})$ where $r(\boldsymbol{\theta}) = (r_1(\boldsymbol{\theta}), \dots, r_2(\boldsymbol{\theta}))^T$, and b is the following diagonal matrix

$$b = \begin{bmatrix} b_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & b_m \end{bmatrix}$$

then we can write

$$R(\boldsymbol{\theta}) = r(\boldsymbol{\theta})^T br(\boldsymbol{\theta}) \tag{3.2.3}$$
$$= f(\boldsymbol{\theta})^T f(\boldsymbol{\theta}). \tag{3.2.4}$$

The minimum of $R(\boldsymbol{\theta})$ is given when the gradient is equal to zero, i.e. when

$$\nabla R(\boldsymbol{\theta}) = 2Df(\boldsymbol{\theta})f(\boldsymbol{\theta}) = \mathbf{0} \tag{3.2.5}$$

where $Df(\boldsymbol{\theta})$ is the Jacobian matrix of $f(\boldsymbol{\theta})$. For convenience denote the Jacobian matrix by $J(\boldsymbol{\theta})$ where

$$J(\boldsymbol{\theta}) = Df(\boldsymbol{\theta}) = \begin{bmatrix} \frac{\partial f_1}{\partial \theta_1} & \cdots & \frac{\partial f_1}{\partial \theta_l} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial \theta_1} & \cdots & \frac{\partial f_m}{\partial \theta_l} \end{bmatrix}$$

Newton's method for nonlinear least squares is an iterative technique that uses a arbitrary search vector to calculate each iteration. The search vector for the k^{th} iteration is be denoted by $\mathbf{L}_k$ and $\boldsymbol{\theta}_k$ is the value of $\boldsymbol{\theta}$ for the k^{th} iteration. Then we have,

$$\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k + \mathbf{L}_k \tag{3.2.6}$$

Expand $\nabla R(\boldsymbol{\theta})$ using a multivariable Taylor series with respect to $\boldsymbol{\theta}$ about $\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k$ we have

$$\nabla R(\boldsymbol{\theta}_{k+1}) = \nabla R(\boldsymbol{\theta}_{k+1}) + D(\nabla R(\boldsymbol{\theta}_k))(\boldsymbol{\theta}_{k+1} - \boldsymbol{\theta}_k) + \ldots \tag{3.2.7}$$
$$= 2Df(\boldsymbol{\theta}_k)^T f(\boldsymbol{\theta}_k) + D(2Df(\boldsymbol{\theta}_k)^T f(\boldsymbol{\theta}_k))(\boldsymbol{\theta}_{k+1} - \boldsymbol{\theta}_k) + \ldots \tag{3.2.8}$$

Ignoring higher order terms, the expansion becomes

$$\nabla R(\boldsymbol{\theta}_k + \mathbf{L}_k) = 2Df(\boldsymbol{\theta}_k)^T f(\boldsymbol{\theta}_k) + 2\left[\sum_{i=1}^{m} f_i(\boldsymbol{\theta}_k) D^2 f_i(\boldsymbol{\theta}_k) + Df(\boldsymbol{\theta}_k)^T Df(\boldsymbol{\theta}_k)\right]\mathbf{L}_k, \tag{3.2.9}$$

where $D^2 f_i(\boldsymbol{\theta})$ is the Hessian matrix of $f_i(\boldsymbol{\theta})$ which is given by,

$$D^2 f_i(\boldsymbol{\theta}) = \begin{bmatrix} \frac{\partial^2 f_i}{\partial \theta_1^2} & \cdots & \frac{\partial^2 f_i}{\partial \theta_1 \partial \theta_l} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_i}{\partial \theta_l \theta_1} & \cdots & \frac{\partial^2 f_i}{\partial \theta_l^2} \end{bmatrix}$$

Let,

$$H(\boldsymbol{\theta}_k) = \sum_{i=1}^{m} f_i(\boldsymbol{\theta}_k) D^2 f_i(\boldsymbol{\theta}_k) \tag{3.2.10}$$

Set equation (3.2.9) to zero, to get the following expression,

$$[H(\boldsymbol{\theta}_k) + J(\boldsymbol{\theta}_k)^T J(\boldsymbol{\theta}_k)]\mathbf{L}_k = -J(\boldsymbol{\theta}_k)^T f(\boldsymbol{\theta}_k), \tag{3.2.11}$$

which can be used to solve for the search vector L_k, The algorithm for Newton's method for nonlinear least squares is given as follows,

---

Require: Initial guess $\boldsymbol{\theta}_0$, Error tolerance $\boldsymbol{\epsilon}$
  for $k = 1, \dots,$ Maximum Iterations **do**
    calculate $H(\boldsymbol{\theta}_k)$, $J(\boldsymbol{\theta}_k)$,$f(\boldsymbol{\theta}_k)$
    solve $[H(\boldsymbol{\theta}_k) + J(\boldsymbol{\theta}_k)^T J(\boldsymbol{\theta}_k)]\mathbf{L}_k = -J(\boldsymbol{\theta}_k)^T f(\boldsymbol{\theta}_k)$
    set $\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k + \mathbf{L}_k$
    if $R(\boldsymbol{\theta}) \leq \boldsymbol{\epsilon}$ **then**
      break
    end if
  end for

In the case of the lottery market prospect we have $m = 7$, $n = 7$, $l = 2$, $\boldsymbol{\theta} = (\alpha, \gamma)^T$, $\mathbf{x}^{(i)} = (x_1^{(i)}, x_2^{(i)}, x_3^{(i)}, x_4, \dots, x_7)$ since x_4 to x_7 are fixed, and for $i = 1, \dots m$,

$$r_i(\boldsymbol{\theta}) = V(\mathbf{x}^{(i)}, \mathbf{p}; \boldsymbol{\theta}) - v(5; \alpha) \tag{3.2.12}$$

$$= w(p_1; \gamma)v(x_1^{(i)}; \alpha) + [w(p_1 + p_2; \gamma) - w(p_1; \gamma)]v(x_2^{(i)}; \alpha) \tag{3.2.13}$$

$$+ \dots + [w(p_1 + \dots + p_7; \gamma) - w(p_1 + \dots + p_6; \gamma)]v(x_7; \alpha) \tag{3.2.14}$$

The disadvantage with this method is that it requires the computation of the first and second derivative of the the residual, but because the value function, $v(x)$, and the decision weighting function, $w(p)$, are both twice differentiable and so this was not a problem. Another disadvantage is that a sufficiently good initial guess is required [25]. Thus, we use the solution obtained from the local search optimization as the initial guess.

3.3 Results

The values of α and γ that minimize the mean square error is summarized below in Table 3.2 with the given mean squared error (MSE):
The value function for $\alpha = 0.2384$ is given below in Figure 3.1 along with the graph of the linear value function $v(x) = x$. The decision weighting function is given in Figure 3.2

$\alpha = 0.2384$	$\gamma = 0.2578$	$MSE = 8.1679 \times 10^{-4}$

Table 3.2: Parameter values that minimize the weighted mean square error.

for $\gamma = 0.2578$ as well as the graph of $w(p) = p$. The graphs of $v(x) = x$ and $w(p) = p$ are given so one can see the extreme differences between the linear case and the estimated parameters.

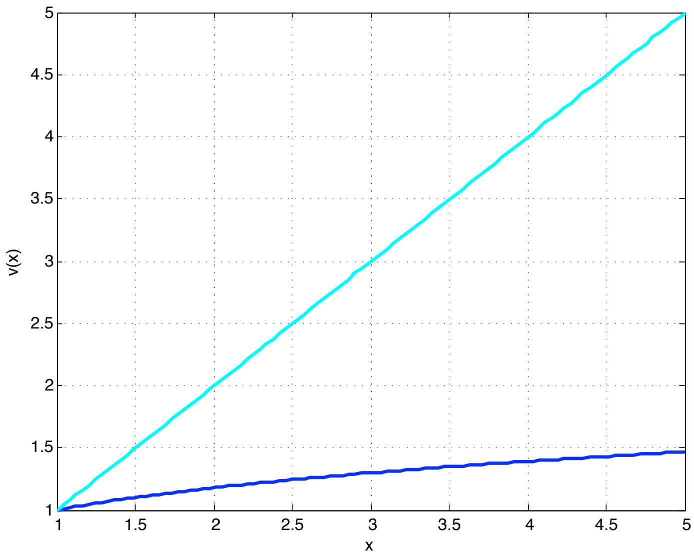

Figure 3.1: The graph of $v(x) = x^{0.2384}$ is given in dark blue, and the graph of $v(x) = x$ is given in light blue

The estimated value function and the estimated decision weighting function are obviously unrealistic. The value functions marginal value diminishes too rapidly as wealth increases and the decision weighting function implies that individuals do not distinguish between probabilities of approximately 0.1 to 0.8, which is obviously not the case in the real world. This kind of decision weighting function is essentially a step function, and corresponds to a decision weighting function that is commonly seen in small children. They distinguish between possible, and impossible events, and the rest are considered maybes. Experts, however, usually exhibit a decision weighting function that is relatively linear when making decisions in their area of expertise [15]. This might actually be how people behave in the lottery market since it is a little difficult to be an expert in playing lottery. However, individuals clearly do not act like this in other areas of decision making, they can distinguish between probabilities of 0.1 to 0.8, thus, these estimates are useless in out of sample prediction, but does provide interesting insight into behaviours in the lottery market.

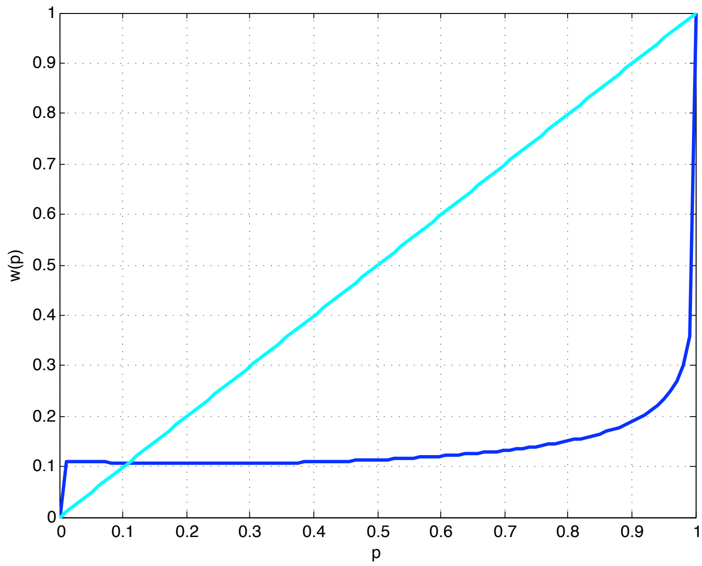

Figure 3.2: The graph of $w(p) = \frac{p^{0.2578}}{(p^{0.2578}+(1-p)^{0.2578})^{\frac{1}{0.2578}}}$ is given in dark blue and the graph of $w(p) = p$ is given in light blue

Another argument is that the parametric form cannot account for an extreme probability prospect like the lottery prospect. The parametric form might not be suitable for the degree of overweighting of probabilities seen in the lottery markets. Even if you look at the approach where the ticket purchaser's initial assets are used as a reference point, the over weighting of small probabilities would need to be even more dramatic to compensate for the loss of the ticket price.

It is interesting to look at how the mean squared error varies with changes in α and γ. Figure 3.3 gives the surface plot of the mean square error for combinations of α and γ varying from 0.45 to 0.6 and 0.55 and 0.8 respectively. This figure suggests that the mean squared error decreases for similar values of α and γ.

If we look at the value of γ that minimizes the weighted mean squared error for fixed values of α then we can we can see that as α increases, γ increases almost linearly. Figure 3.4 displays this result. The mean squared error, however, increases as α increases, except for small α. This is shown in Figure 3.5. As α increases $v(x)$ also increases. So to compensate for this increase, γ increases, which causes the overweighting of small probabilities to become less pronounced.

The next section will look at the feasibility region of the prospect and the discuss reasons why the estimated parameters may be realistic.

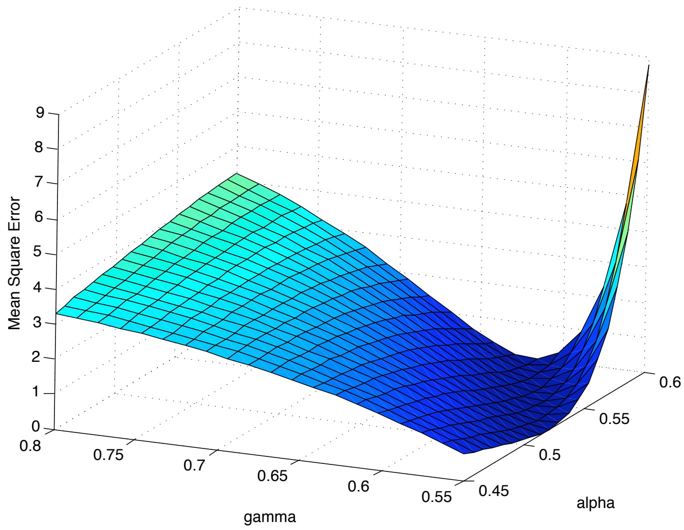

Figure 3.3: The surface plot of α, γ, and the mean square error

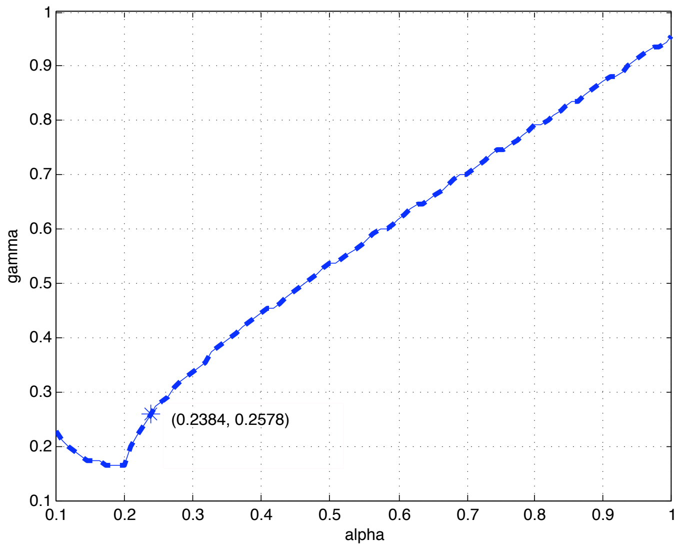

Figure 3.4: The value of γ that minimizes the MSE for fixed α.

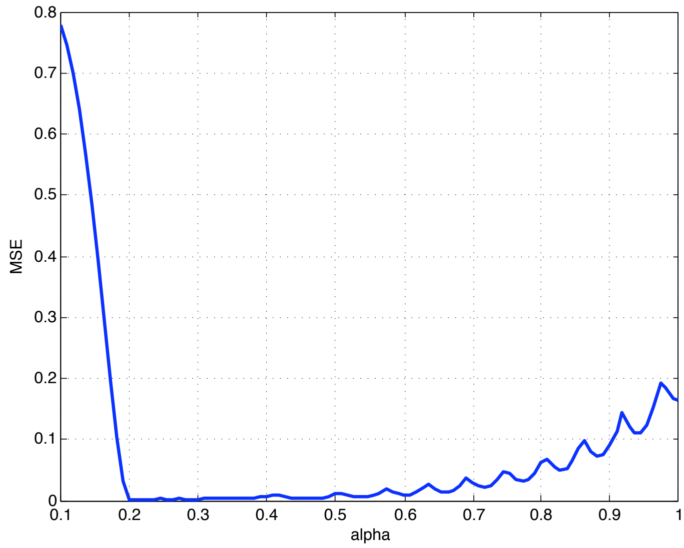

Figure 3.5: The MSE for fixed α.

3.4 Discussion

The parameter estimates that minimize the weighted mean squared error given in Table 3.2 is not the only parameter combination that can account for the behaviour in the lottery market. Any parameter combination that satisfies

$$V(\mathbf{x}, \mathbf{p}) \geq v(5), \tag{3.4.1}$$

is a feasible set of parameters, and those individuals with parameters in the feasibility set would be willing to purchase a lottery ticket. Figure 3.6 shows the approximate feasibility curve, which was calculated using average data, and by holding α constant for values of α between $[0.1, 1]$. Any parameter combination above the feasibility line does not satisfy (3.4.1), and anything below the line does. Thus, parameter combinations below the line are considered feasible. Table 3.3 shows the parameters estimated by Kahneman and Tversky [31], Camerer and Ho [7], and Gonzalez and Wu [15]. The parameter combinations are also displayed in Figure 3.6.

As was seen from Figure 3.6 and Table 3.3 the parameters estimated from Camerer and Ho, and Gonzalez and Wu are not feasible. If an individual exhibited these preferences consistent with their estimations then they would not purchase the lottery ticket. Kahneman and Tversky's estimation gave a value of the prospect, $V(\mathbf{x}, \mathbf{p})$, that is quite a lot larger then the value of the price of the lottery ticket. Therefore, an individual that exhibited preferences consistent with their parameters would be willing to purchase a lottery ticket. This would also suggest that if the Kahneman and Tversky's parameter estimates were

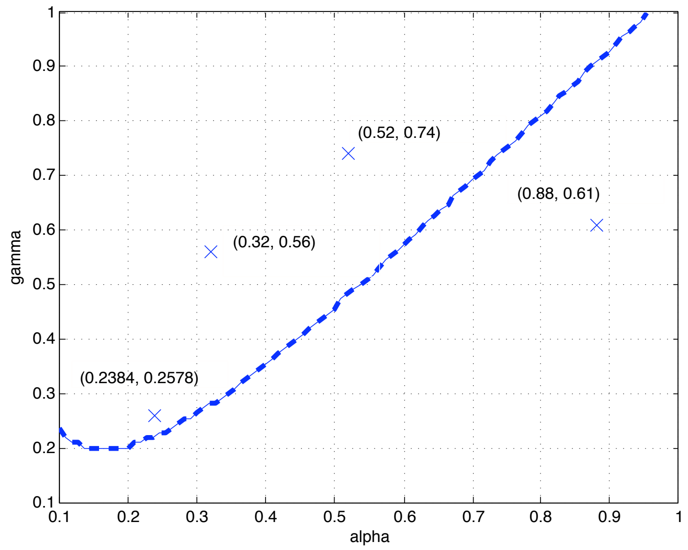

Figure 3.6: The feasibility region for the lottery prospect is given below the line

Source	α	γ	$V(\mathbf{x}, \mathbf{p}) - v(5)$
Kahneman & Tversky	0.88	0.61	114.2346
Camerer & Ho	0.32	0.56	-1.2414
Gonzalez & Wu	0.52	0.74	-1.8034
This paper	0.2384	0.2578	-5.6350×10^{-5}

Table 3.3: Parameter estimates for α, and γ and their feasibility

truly ones consistent with individuals who buy the lottery tickets then the lottery company could increase the price of the ticket without decreasing sales. However, instead one could conclude that Kahneman and Tversky's estimations are not correct and according to the lottery market presented here and by looking at Figure 3.6 we could say that either the value of α is too high, the value of γ is too low, or a combination of both.

To explore the feasibility of parameter combinations see Neilson and Stowe [21]. They have explored the appropriateness of the combinations of the parameterized components together. They found that gambling on small probability outcomes and other kinds of behaviour could not be accounted for with most parameter combinations using Kahneman and Tversky's parametric form.

There is another way of viewing the parameter estimates in Table 3.2 that makes the estimates seem more realistic. Cook and Clotfelter [9] did a study on lottery markets. They found that individuals judge the probability of winning the lottery based on the frequency of winners. This means that lotteries that have greater sales can get away with having a lower probability of winning than other lotteries, which is consistent with lotteries in the real world. Larger state lotteries offer a game at decreased odds but with the same perceived probability of winning as the smaller state [9]. This is most likely because the probability of winning the lottery is never explicitly publicised but what is publicised is the size of the winnings and the number of winners.

If we assume that people judge probabilities based on the frequency of winners then we can consider this prospect under uncertainty. This means that people most likely employ the two-stage decision process where they first make a judgment about the probabilities then that probability gets transformed by the decision weighting function under risk. There also exists evidence that an individual's judged probabilities exhibit bounded subadditivity. Fox *et al* [13] finds that option traders overweight perceived probabilities that are low and underweight perceived probabilities of moderate to high. Others have also found evidence of this phenomenon such as Kilka and Weber [19], and Fox and Tversky [28]. This suggests that the parameter estimates would be more realistic if we take into account bounded subadditivity of the judged probabilities for the lottery prospect. Thus, the parametric form suggested by Kahneman and Tversky may in fact be able to account for the lottery market if we consider the lottery prospect as a decision under uncertainty.

3.5 Conclusions

Cumulative prospect theory takes many attributes into account: reference dependence, diminishing sensitivity, nonlinear preferences, loss aversion, risk seeking, and source dependence. It has been used to explain the equity-premium puzzle, the disposition effect, labour supply, racetrack betting, and so on [6]. It has also been applied to asset allocation models, portfolio optimization, asset pricing, medical decision making, and even political decision making. Cumulative prospect theory generalises expected utility theory into a model that can account for a diverse range of risk attitudes, and modifies the original prospect theory to satisfy stochastic dominance.

The purpose of this paper was to investigate and analyze the cumulative prospect theory model and use data from a real world context as opposed to data from a controlled experiment to estimate the parameters from the functional form proposed by Kahneman and Tversky [31]. The estimates predicted that individuals exhibited a decision weighting function that was essentially a step function and a value function that displayed rapid diminishing sensitivity. These parameters are clearly not realistic, however, they may be explained by two reasons. The first is that the price of the lottery ticket may not be the true market price, and that the lottery corporation could increase the price of the ticket without suffering loss of sales. The second is that individuals who buy the ticket are making a decision under uncertainty, where the they exhibit a two-stage process; they over estimate the judged probability and then again overweight the probability. Both explanations need more research to conclude for certain. Further improvements to the analysis are as follows: use insurance data to estimate the parameters in the loss domain, look at the case where the lottery ticket is considered a loss, look at other parametric forms, consider parameter-free elicitation, Include data from other lotteries, and account for judgments in perceived probabilities

Appendix A

Matlab Code

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%Local Search Algorithm:
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%function [Weight, MinError, gamma, alpha] = CPT(Alpha, Gamma)
clear

p1 = 1/13983816;
p2 = 1/2330636;
p3 = 3/166473;
p4 = 15/332948;
p5 = 12915/13983816;
p6 = 4305/3495954;
p7 = 4100/249711;

p = [p1; p1+p2; p1+p2+p3; p1+p2+p3+p4;
    p1+p2+p3+p4+p5; p1+p2+p3+p4+p5+p6;
    p1+p2+p3+p4+p5+p6+p7;];

x = [12868476.80, 594754.65, 34371.85, 4800, 320, 160, 20;
    14526918.33, 731742.61, 35359.35, 4800, 320, 160, 20;
    13203704.00, 462040.22, 33144.89, 4800, 320, 160, 20;
    14136913.94, 643626.97, 34419.09, 4800, 320, 160, 20;
    12794725.56, 499868.89, 32461.11, 4800, 320, 160, 20;
    24845828.57, 686297.86, 41650.71, 4800, 320, 160, 20;
    17485475.00, 948558.57, 43402.86, 4800, 320, 160, 20];
```

```
L = 5;

Alpha = linspace(0.45, 0.6, 20);
Gamma = linspace(0.55, 0.8, 20);

Weight = 1;

b = [100; 69; 45; 33; 18; 7; 7]./279;
B  = b*ones(1, length(Gamma));
C = L.*ones(length(b), 1);
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%Calculating decision weights using Kahneman & Tversky's function
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
if Weight == 1
    w = (p.^Gamma(1))./((p.^Gamma(1)
    + (1 - p).^Gamma(1)).^(1/Gamma(1)));

    for i =  2: length(Gamma)
        gamma = Gamma(i);
        ww = (p.^gamma)./((p.^gamma + (1 - p).^gamma).^(1/gamma));
        w = [w, ww];
    end
    w = w';
    W = [zeros(length(Gamma), 1), w(1:length(Gamma), 1:length(p)-1)];
    w = w - W;
end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%Calculating decision weights using Prelec's function
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
if Weight == 2
    w = exp(-(-log(p)).^Gamma(1));

    for i =  2: length(Gamma)
        gamma = Gamma(i);
        ww = exp(-(-log(p)).^Gamma(i));
        w = [w, ww];
```

```
    end
    w = w';
    W = [zeros(length(Gamma), 1), w(1:length(Gamma), 1:length(p)-1)];
    w = w - W;
end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%Calculating the value function, the expected value, V, and the MSE
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
for i = 1:length(Alpha)
    alpha = Alpha(i);
    v = x.^alpha;
    V = v * w';
    CC = (C.^alpha) * ones(1, length(Gamma));
    if i == 1
        Error = sum(B.*((V - CC).^2));
    else
        Error = [Error; sum(B.*((V - CC).^2))];
    end
end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%Finding the values of gamma and alpha that minimize the MSE
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
Error = Error;
[MinError1, I1] = min(Error);
[MinError2, I2] = min(MinError1);
MinError = Error(I1(I2), I2)
gamma = Gamma(I2)
alpha = Alpha(I1(I2))

surf(Alpha, Gamma, Error')
xlabel('alpha')
ylabel('gamma')
zlabel('Mean Square Error')
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%Nonlinear Least Square
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
clear

iterations = 10000;
epsilon = [10^-5; 10^-5];

%Initial Guess

theta = [0.2376;0.2565];  % theta = [alpha; gamma]

for i = 1:iterations
    if weight == 1
        [J, H, y] = CPT_function(theta);
    end
    theta1 = theta;
    L = (J'*J + H)\(-J'*y);
    theta = theta + L;
    [E] = Error(theta(1), theta(2));

    if  E <= epsilon(1) & abs(theta1 - theta)<= epsilon(2)
        break
    end
end

theta
E
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
function[A, S, y] = CPT_function(theta)

alpha = theta(1);
gamma = theta(2);

x = [12868476.80, 594754.65, 34371.85, 4800, 320, 160, 20;
    14526918.33, 731742.61, 35359.35, 4800, 320, 160, 20;
    13203704.00, 462040.22, 33144.89, 4800, 320, 160, 20;
    14136913.94, 643626.97, 34419.09, 4800, 320, 160, 20;
    12794725.56, 499868.89, 32461.11, 4800, 320, 160, 20;
    24845828.57, 686297.86, 41650.71, 4800, 320, 160, 20;
    17485475.00, 948558.57, 43402.86, 4800, 320, 160, 20];

b = [100; 69; 45; 33; 18; 7; 7]./279;

C = 5*ones(length(b), 1);

p1 = 1/13983816;
p2 = 1/2330636;
p3 = 3/166473;
p4 = 15/332948;
p5 = 12915/13983816;
p6 = 4305/3495954;
p7 = 4100/249711;

p = [p1; p1+p2; p1+p2+p3; p1+p2+p3+p4;
    p1+p2+p3+p4+p5; p1+p2+p3+p4+p5+p6;
    p1+p2+p3+p4+p5+p6+p7;];

b = sqrt([100; 69; 45; 33; 18; 7; 7]./279);
b = diag(b);
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
w = (p.^gamma)./((p.^gamma + (1 - p).^gamma).^(1/gamma));
W = [0; w(1:length(w)-1)];
w = w - W;
```

```
v = x.^alpha;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Calulation of derivative and second of y with respect to gamma
f1 = p.^gamma + (1 - p).^gamma;
f2 = (p.^gamma).*log(p) + ((1-p).^gamma).*log(1-p);
f3 = -log(f1)./(gamma.^2) + f2./(gamma*f1);

w1 = ((p.^gamma)./f1.^(1/gamma)).*(log(p) + log(f1)./(gamma.^2)
- f2./(gamma*f1));

w2 = log(p).^2 - 2*log(p).*f3 + f3.^2 - 2*f3./gamma;
w2 = w2 + (f2.^2)./(gamma.*f1.^2) + ((p.^gamma).*log(p)
+ ((1-p).^gamma).*log(1-p))./(gamma.*f1);
w2 = ((p.^gamma)./f1.^(1/gamma)).*w2;

W1 = [0; w1(1:length(w)-1)];
W2 = [0; w2(1:length(w)-1)];

w1 = w1 - W1;
w2 = w2 - W2;

y_gamma = v*w1;
y_gamma2 = v*w2;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Calulation of derivative & second derivative of y w.r.t. alpha
v1 = (x.^alpha).*log(x);
v2 = (x.^alpha).*(log(x).^2);

y_alpha = v1*w - (C.^alpha).*log(C);
y_alpha2 = v2*w - (C.^alpha).*(log(C).^2);
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Second derviative of y w.r.t. alpha and gamma
y_alpha_gamma = v1*w1;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% The Jacobian Matrix
```

```
A = [b*y_alpha, b*y_gamma];
% The Hessian Matrix
S = [sum(b*y_alpha2), sum(b*y_alpha_gamma);
       sum(b*y_alpha_gamma), sum(b*y_gamma2)];
% The weighted residuals
y = b*(v*w - C.^alpha);
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%Calculates the weighted MSE given alpha and gamma
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
function [E] = Error(alpha, gamma)

p1 = 1/13983816;
p2 = 1/2330636;
p3 = 3/166473;
p4 = 15/332948;
p5 = 12915/13983816;
p6 = 4305/3495954;
p7 = 4100/249711;

p = [p1; p1+p2; p1+p2+p3; p1+p2+p3+p4;
    p1+p2+p3+p4+p5; p1+p2+p3+p4+p5+p6;
    p1+p2+p3+p4+p5+p6+p7;];

x = [12868476.80, 594754.65, 34371.85, 4800, 320, 160, 20;
    14526918.33, 731742.61, 35359.35, 4800, 320, 160, 20;
    13203704.00, 462040.22, 33144.89, 4800, 320, 160, 20;
    14136913.94, 643626.97, 34419.09, 4800, 320, 160, 20;
    12794725.56, 499868.89, 32461.11, 4800, 320, 160, 20;
    24845828.57, 686297.86, 41650.71, 4800, 320, 160, 20;
    17485475.00, 948558.57, 43402.86, 4800, 320, 160, 20];

C = (5^alpha)*ones(7, 1);
b = [100; 69; 45; 33; 18; 7; 7]./279;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
w = (p.^gamma)./((p.^gamma + (1 - p).^gamma).^(1/gamma));
w = w;
W = [0; w(1:length(p)-1)];
w = w - W;
v = x.^alpha;
V = v * w;

E = sum(b.*((V - C).^2));
```

References

[1] Mark 6, official website. http://bet.hkjc.com/marksix/, June 2008.

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

[3] Richard Aster, Brian Borchers, and Clifford Thurber. *Parameter Estimation and Inverse Problems*. Elsevier Academic Press, 2005.

[4] Manel Baucells and Franz H. Heukamp. Stochastic dominance and cumulative prospect theory. *Management Science*, 52(9):1409–1423, 2006.

[5] Han Bleichrodt and Jose Luis Pinto. A parameter-free elicitation of the probability weighting function in medical decision analysis. *Management Science*, 46(11):1485–1496, 2000.

[6] Colin F. Camerer. *Prospect Theory in the Wild: Evidence from the Field*. Cambridge University Press, 2000.

[7] Colin F. Camerer and Teck-Hua Ho. Violations of the betweenness axiom and nonlinearity in probability. *Journal of Risk and Uncertainty*, 8:167–196, 1994.

[8] Alain Chateauneuf and Peter Wakker. An axiomatization of cumulative prospect theory for decision under risk. *Journal of Risk and Uncertainty*, 18(2):137–145, 1999.

[9] Philip J. Cook and Charles T. Clotfelter. The peculiar scale economies of lotto. *The American Economic Review*, 83(3):634–643, 1993.

[10] Stephen J. Dubner. Calculating the irrational in economics. The New York Times, June 2003.

[11] Hein Fennema and Peter Wakker. Original and cumulative prospect theory: A discussion of empirical differences. *Journal of Behavioral Decision Making*, 10:53–64, 1997.

[12] Peter C. Fishburn. Retrospective on the utility theory of von neumann and morgenstern. *Journal of Risk and Uncertainty*, 2:127–158, 1989.

[13] Craig R. Fox, Brett A. Rogers, and Amos Tversky. Option traders exhibit subadditive decision weights. *Journal of Risk and Uncertainty*, 13(1):5–17, 1996.

[14] Milton Friedman and L. J. Savage. The utility analysis of choices involving risk. *The Journal of Political Economy*, 56(4):279–304, 1948.

[15] Richard Gonzalez and George Wu. On the shape of the probability weighting function. *Cognitive Psychology*, 38:129–166, 1999.

[16] Daniel Kahneman, Jack L. Knetsch, and Richard H. Thaler. The endowment effect, loss aversion, and status quo bias. *Journal of Economic Perspectives*, 5(1):193–206, 1991.

[17] Daniel Kahneman and Amos Tversky. Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–292, 1979.

[18] Daniel Kahneman and Amos Tversky. Choices, values, and frames. *American Psychologist*, 39:341–350, 1984.

[19] Michael Kilka and Martin Weber. What determines the shape of the probability weighting function under uncertainty. *Management Science*, 47(12):1712–1726, 2001.

[20] Harry Markowitz. The utility of wealth. *The Journal of Political Economy*, 60(2):151–158, 1952.

[21] William Neilson and Jill Stowe. A further examination of cumulative prospect theory parameterizations. *The Journal of Risk and Uncertainty*, 24(1):31–46, 2002.

[22] Drazen Prelec. The probability weighting function. *Econometrica*, 66(3):497–527, 1998.

[23] John Quiggin. A theory of anticipated utility. *Journal of Economic Behavior and Organization*, 3:323–343, 1982.

[24] Matthew Rabin and Richard H. Thaler. Anomalies: Risk aversion. *Journal of Economic Perspectives*, 15(1):219–232, 2001.

[25] L. E. Scales. *Introduction to Nonlinear Optimization*. MacMillan Publishers Ltd., 1985.

[26] Jeremy J. Siegal and Richard H. Thaler. Anomalies: The equity premium puzzle. *The Journal of Economic Perspectives*, 11(1):191–200, 1997.

[27] Henry P. Stott. Cumulative prospect theory's functional menagerie. *Journal of Risk and Uncertainty*, 32:101–130, 2006.

[28] Amos Tversky and Craig R. Fox. Weighing risk and uncertainty. *Psychological Review*, 102(2):269–283, 1995.

[29] Amos Tversky and Chip Heath. Preference and belief: Ambiguity and competence in choice under uncertainty. *Journal of Risk and Uncertainty*, 4(1):5–28, 1991.

[30] Amos Tversky and Daniel Kahneman. Loss aversion in riskless choice: A reference-dependent model. *The Quarterly Journal of Economics*, 106(4):1039–1061, 1991.

[31] Amos Tversky and Daniel Kahneman. Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5:297–323, 1992.

[32] Amos Tversky and Peter Wakker. Risk attitudes and decision weights. *Econometrica*, 63(6):1255–1280, 1995.

[33] Peter Wakker and Daniel Deneffe. Eliciting von neumann-morgenstern utilities when probabilities are distorted or unknown. *Management Science*, 42(8):1131–1150, 1996.

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

[35] George Wu and Richard Gonzalez. Curvature of the probability weighting function. *Management Science*, 42(12):1676–1690, 1996.

[36] George Wu and Richard Gonzalez. Nonlinear decision weights in choice under uncertainty. *Management Science*, 45(1):74–85, 1999.