A Midpoint Technique for Easily Measuring Prospect Theory's Probability Weighting*

GIJS VAN DE KUILEN[a], PETER P. WAKKER[b], & LIANG ZOU[a]

August, 2006

a: Dept. of Economics, University of Amsterdam, Roetersstraat 11, Amsterdam, 1018 WB, The Netherlands;

b: Econometric Institute, Erasmus University, P.O. Box 1738, Rotterdam, 3000 DR, The Netherlands, --31-(0)10 – 408.12.65 (O), --31-(0)10 – 408.12.78/12.64 (S), --31-(0)10 – 408.91.62 (F), Wakker@few.eur.nl

ABSTRACT. Prospect theory can better describe risky choices than classical expected utility because it makes the plausible assumption that risk aversion is driven as much by the way people feel about probabilities (probability weighting) as about outcomes (utility). This leads to better predictions but, as a price to pay, probability weighting is more difficult to measure than utility, which may explain why many economists today continue to use expected utility. This paper mitigates the drawback mentioned by introducing a new method for measuring probability weighting that is simpler and more efficient than methods used before. The new method is implemented in an experiment. Most participants exhibited a convex weighting function, implying pessimism and enhancing risk aversion. This finding supports recent claims that utility is less concave than was traditionally thought in studies that ascribed all risk aversion to concave utility.

JEL-Classification: D81, C60, C91

KEYWORDS: Prospect Theory; Probability Weighting; Pessimism

* Han Bleichrodt made useful comments.

1. INTRODUCTION

Many empirical studies have demonstrated that prospect theory of Kahneman and Tversky yields better predictions about the way in which people take risks than classical expected utility. Nevertheless the majority of papers in economics today still use the classical expected utility model to analyze risky decisions. One reason for the slow acceptance of prospect theory may be that this theory, in being more general than expected utility, is also more difficult to apply. It takes more work to derive predictions from theoretical analyses and from empirical measurements of the new components of risk attitude. This paper will contribute to reducing the second problem, by making empirical measurements easier than they were before.

Prospect theory introduces two new concepts that are not present in expected utility. The first is loss aversion, which entails that people take reference points and weigh outcomes below the reference point (losses) more heavily than gains. We will confine our attention to positive outcomes (gains), so that loss aversion plays no role in this paper. The second new concept of prospect theory concerns probability weighting and this is the topic of our investigation. Whereas classical economics ascribes risk aversion solely to utility, being a nonlinear scale for the evaluation of outcomes, it is highly plausible that risk aversion be driven as much by a nonlinear scale for the evaluation of probabilities. This is what probability weighting captures. To illustrate the plausibility of probability weighting as a factor to explain risky behavior, the coexistence of gambling and insurance, known as a paradox for classical theories that explain risk attitude solely in terms of utility, can readily be explained through probability weighting (Tversky & Kahneman 1992). Our paper will make probability weighting easy to measure. Thus, it becomes easier to use this important component of risk attitude to improve predictions about the risky behavior of people.

Our method is based on a technique for measuring midpoints of decision weights. For simplicity of presentation, we will present our new technique in an experiment for decision under risk (known probabilities). As we will demonstrate in the theoretical analysis, our measurement method can equally well be used to measure nonadditive weighting functions for uncertainty with unknown probabilities (Gilboa 1987; Schmeidler 1989; Tversky & Kahneman 1992). Hence, it can also be used to examine ambiguity attitudes.

Besides of interest on its own, the measurement of probability weighting is also important for the measurement of utility. Classical utility measurements in economics and other domains (Dybvig & Polemarchakis 1981; Gold et al. 1996, Keeney & Raiffa 1976) invariably assumed expected utility, and ascribed all risk aversion to concavity of utility. The index of (relative) risk aversion, which only captures concavity of utility, is generally used as index of risk aversion indeed, as even expressed by its name. If, however, risk aversion is partly generated by probability weighting, then utility is less concave than suggested by the classical measurements, and all predictions and prescriptions based on the utility component are affected (Rabin 2000; Young 1990). This constitutes one of the major challenges to the foundations of economics today.

We use our method in an experiment with N=78 participants. In the literature, the most common finding is a combination of inverse-S shaped and convex-shaped curves, where the inverse-S shape has been found to be prevailing in most studies (Abdellaoui 2000; Bleichrodt & Pinto 2000; Gonzalez & Wu 1999; Tversky & Fox 1995; Tversky & Kahneman 1992). Convexity was assumed for instance in Starmer (1992, p. 818 and pp. 819-820, expressed as concavity of the dual of w), and convex shapes were found to be prevailing in Jullien and Salanié (2000). In the present study we find both shapes too, with convex shapes prevailing.

The remainder of this paper is organized as follows. Section 2 briefly presents Tversky & Kahneman's (1992) prospect theory, and presents details of techniques for measuring probability weighting used before in the literature. Section 3 introduces our new measurement technique, first for risk, and then, at the end, for uncertainty. The experiment is in Section 4, with results in Section 5, and a discussion in Section 6. Section 7 concludes.

2. PROSPECT THEORY AND EARLIER MEASUREMENT TECHNIQUES

Outcomes are monetary, with $\mathbb{R}^+$ the outcome set. We do not consider negative outcomes (losses). A *prospect* is a probability distribution over outcomes taking only finitely many outcomes; $(p_1: x_1,\ldots, p_n: x_n)$ denotes a prospect yielding outcome x_i with probability p_i, i = 1,…, n, where probabilities are nonnegative and sum to one. For two outcomes we often suppress the second probability, and write (p:x, y) for (p:x, 1–p:y). U: $\mathbb{R}^+ \to \mathbb{R}$ denotes the utility function. It is continuous and strictly increasing. Probabilities are transformed into probability weights through a strictly increasing and continuous *probability weighting function* $w:[0,1]\to[0,1]$, with w(0) = 0 and w(1) = 1.

In this paper, *prospect theory* refers to the modern version of prospect theory, introduced by Tversky & Kahneman (1992). It corrects some theoretical problems of the original 1979 version, and can also deal with uncertainty, i.e. the case of unknown probabilities. Under prospect theory, each outcome x_i is weighted by a subjective *decision weight* denoted π_i. The decision weight is obtained by subtracting the weighted probability of receiving only an outcome rank-ordered strictly better than x_i from the weighted probability of receiving x_i or an outcome rank-ordered strictly better. That is:

$$\pi_i = w(p_i + \cdots + p_1) - w(p_{i-1} + \cdots + p_1), \tag{2.1}$$

for a prospect $(p_1{:}x_1,\ldots, p_n{:}x_n)$ with $x_1 \geq \cdots \geq x_n$. The decision weight of x_i is the marginal w contribution of p_i to the probability of receiving better outcomes. The evaluation of the prospect is

$$\sum_{i=1}^{n} \pi_i U(x_i)\,. \tag{2.2}$$

Expected utility results if w(p) = p for all p, so that $\pi_i = (p_i + \cdots + p_1) - (p_{i-1} + \cdots + p_1) = p_i$ for all i.

Most measurements in the literature have used parametric fittings. Then a series of direct choices is elicited and the utility function and the subjective probability weighting function are estimated jointly from the data. A drawback of parametric fittings is that if the assumed families differ from the true underlying functional form, then conclusions based on these fittings need not be valid. For example, several parametric fittings considered weighting functions that are only globally convex or globally concave (Hey & Orme 1994), or only inverse-S (Donkers, Meelenberg & van Soest 2001), so that no insight resulted about the prevalence of such shapes relative to other shapes. Our experiment will illustrate this difficulty of parametric fitting.

A second drawback of parametric fitting concerns the joint fitting of utility and probability weighting. The parameter estimates of these functions are interdependent: an overestimation of risk aversion in one component leads to an underestimation in the other, and vice versa. Therefore, Tversky & Kahneman (1981) suggested that "the simultaneous measurement of values and decision weights involves serious experimental and statistical difficulties (p.454)."

Gonzalez & Wu (1999) did not commit to a parametric family but still used fitting techniques that minimize squared distances, based on a complex numerical system that requires much data per participant. In return, their results are very reliable. Abdellaoui (2000) and Bleichrodt & Pinto (2000) provided two more tractable methods for estimating probability weighting functions nonparametrically. Details are given below. As with all other measurements used so far, these methods need a detailed measurement of utility before probability weighting can be measured. The present paper introduces a new and simpler technique that, for each pair of probabilities, can easily infer the midpoint probability in terms of probability weights. It then becomes easy to measure the probability weighting function to any degree of precision. Starting from w(0) = 0 and w(1) = 1, we can measure the p with w(p) = 1/2, i.e. $w^{-1}(1/2)$. Then, with w(0) = 0 and $w^{-1}(1/2)$ available, we can measure $w^{-1}(1/4)$, and, similarly, $w^{-1}(3/4)$, $w^{-1}(1/8)$; etc. We can continue measuring midpoints until we have estimated the curve of w as accurately as we want.

Our method can be used both for parametric and for nonparametric measurements. It provides the first measurement of probability weighting in the literature that does not need a detailed measurement of utility and, hence, provides the most efficient way to measure probability weighting that is presently available. From n observed indifferences we obtain n–2 data points of the probability function (plus 1 data point of utility), whereas Abdellaoui (2000) for instance would obtain only (n–1)/2 data points of probability weighting (plus (n–1)/2 data points of utility). Details of the methods just discussed, preparing for our method, are as follows.

Wakker & Deneffe (1996) proposed to measure utility through indifferences

$$(p: x_{i+1}, y) \sim (p: x_i, Y), \text{ for } x_{i+1} > x_i > Y > y, \quad i = 0,\ldots, n. \tag{2.3}$$

With $\pi = w(p)$, these indifferences imply that $\pi\big(U(x_{i+1}) - U(x_i)\big) = (1-\pi)\big(U(Y) - U(y)\big)$ for all i, so that

$$U(x_{i+1}) - U(x_i) \text{ is the same for all i.}$$

These equalities hold true irrespective of what w(p) is, so that we need not measure a person's subjective w function to make this inference about the person's subjective utility function.

Abdellaoui (2000) proposed to measure probability weighting by first measuring $x_0,\ldots, x_6$ as above, with utility normalized at $U(x_0) = 0$ and $U(x_6) = 1$ so that $U(x_j) = j/6$ for all j. He then elicited $p_1,\ldots, p_5$ such that

$$x_i \sim (p_i{:}\ x_6, x_0) \text{ for all i.}$$

We get $U(x_i) = w(p_i)U(x_6) + (1-w(p_i))U(x_0)$, or $i/6 = w(p_i)1 + (1-w(p_i))0$, so that

$$w(p_i) = i/6 \text{ for all i.}$$

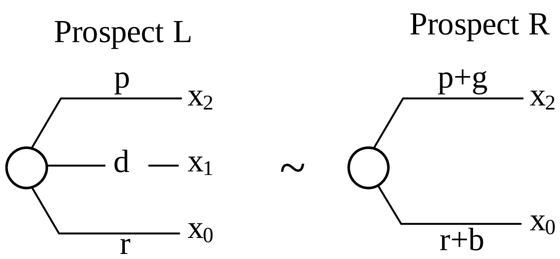
Bleichrodt & Pinto (2000) independently introduced a very similar method, and Etchart (2004) used Abdellaoui's method to measure the weighting function for losses.

These methods make extensive use of the outcome domain and need to carry out a utility measurement at least as detailed as the probability weighting measurement that is desired. For example, to measure the probability with weight 1/32, we need to elicit 32 values x_i, or rely on parametric interpolations. Our method will avoid these complications. Like the methods described, it uses information obtained from utility measurements to elicit decision weights, but it does so in a different and more efficient manner. Blavatskyy (2006) described the general procedure to start with measurements in one dimension, then use this to obtain measurements in the other dimension, possibly using these again to obtain measurements in the first dimension, and so forth. He examined general efficiency principles of such general procedures.

3. A NEW MIDPOINT TECHNIQUE

Our method of measuring midpoints of a weighting function starts with measuring a midpoint of utility. To this end, we measure x_0, x_1, x_2 as in Eq. 2.3, after which x_1 results as the midpoint between x_0 and x_2 in utility units. These x-values will be used throughout what follows. Alternative methods for endogenously deriving utility midpoints from preferences by Ghirardato et al. (2003) and Vind (2003) will be discussed in Section 6. To elicit weighting functions, we elicit indifferences between the prospects $(p{:}x_2, d{:}x_1, r{:}x_0)$ and $(p+g{:}x_2, r+b{:}x_0)$ depicted in Figure 1, with r the residual probability 1–p–d. Here d is the probability mass of x_1 to be divided over the other outcomes. g is the probability mass taken from d and moved to the good outcome x_2, and the remainder b = d–g is moved to the bad outcome x_0.

FIGURE 1. Distributing d's weight evenly over the upper and lower branch



The intuition behind our technique is as follows, stated informally. Figure 2 will illustrate the decision weights derived hereafter. Prospect R in Figure 1 results from prospect L by moving some of the d-probability mass from outcome x_1 up to x_2 and the remaining probability mass down to x_0. The improvement $U(x_2) - U(x_1)$ and the worsening $U(x_1) - U(x_0)$ are equally big. Hence, to preserve indifference, as much decision weight must have been moved up $\big(w(g+p) - w(p)\big)$ as down $\Big((1 - w(1-r-b) - \big(1 - w(1-r)\big) = w(d+p) - w(g+p)\Big)$. From $w(d+p) - w(g+p) = w(g+p) - w(p)$ it follows that $w(g+p)$ must be the midpoint between $w(d+p)$ and $w(p)$. We next state the result formally. Because its proof may be instructive, we give it in the main text.

Theorem 1. The indifference in Figure 1 implies that

$$w(g + p) = \frac{w(p) + w(d+p)}{2}$$

whenever $U(x_2) - U(x_1) = U(x_1) - U(x_0) > 0$.

Proof. We compare the prospect-theory values of the two prospects with $U(x_0)$ subtracted:

$$w(p)\big(U(x_2)-U(x_0)\big) + \big(w(d+p) - w(p)\big)\big(U(x_1) - U(x_0)\big) \;=\; w(g+p)\big(U(x_2) - U(x_0)\big).$$

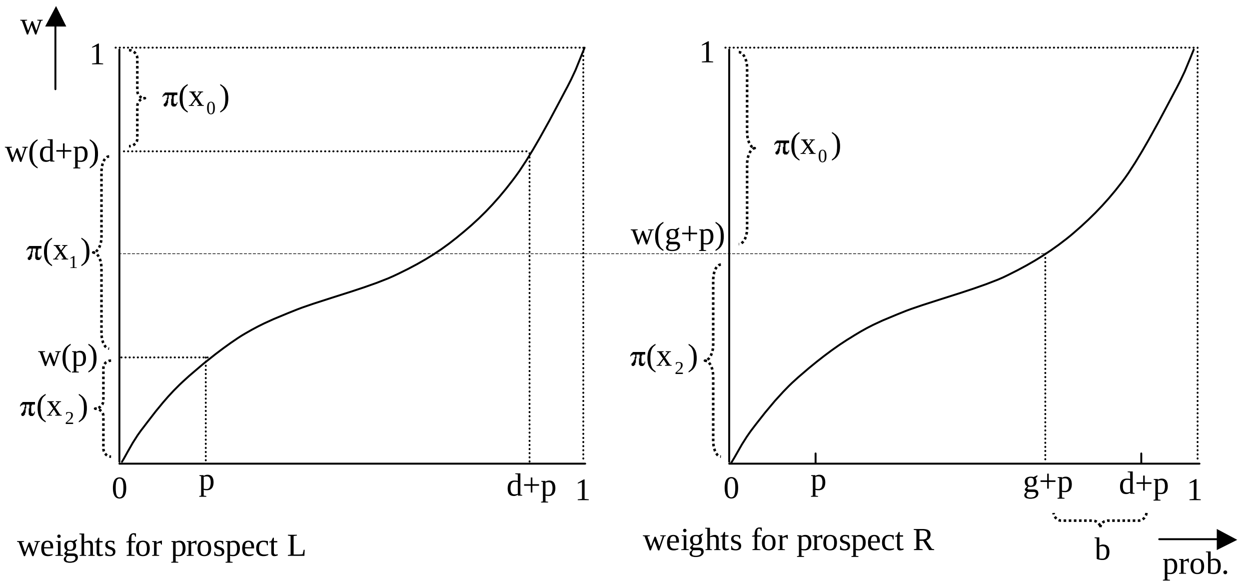
Dividing by $U(x_1) - U(x_0) = \big(U(x_2) - U(x_0)/2\big)$ yields

$$2w(p) + \big(w(d+p)-w(p)\big) \;=\; 2w(g+p), \text{ and}$$

$$\frac{w(p) + w(d+p)}{2} \;=\; w(g+p)$$

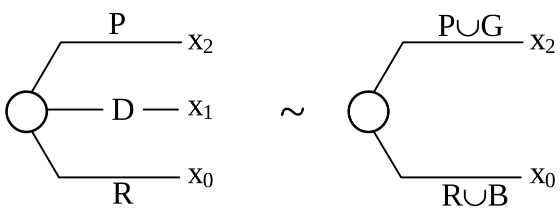
follows, as required. □

FIGURE 2. Decision Weights under Prospect Theory for prospects depicted in Figure 1



Our measurement technique is general in the sense that the weight-midpoint between any two probabilities can be measured directly. The only richness of outcomes needed is that there are three outcomes as above, i.e. equally spaced in utility units.

FIGURE 3. Distributing D's weight evenly over the upper and lower branch



Our technique can readily be extended to Schmeidler's (1989) Choquet expected utility, and prospect theory, for the case of events with unknown probabilities. Figure 3 results from Figure 1 by replacing the probabilities p, d, and r by exhaustive and mutually exclusive events P, D, and R, replacing probability g by a subevent G of R, and replacing probability b by event B = D–G. The indifference in Figure 3 implies that the decision weight of event G captures half the decision weight of event D, so that event G∪P is the weight-midpoint between events P and D∪P, similarly as above. Such a midpoint event G exists for all events P and D∪P if the event space is a continuum or at least is sufficiently rich, as for instance in Gilboa's (1987) preference foundation.

Our paper only considers rank-dependent utility and prospect theory for risk and uncertainty. Other deviations from expected utility include the betweenness models by Chew (1983), Dekel (1986), Gul (1991), and Chew & Tan (2005) for risk, and regret theory (Bell 1982; Loomes & Sugden 1982) and multiple priors (Chateauneuf 1991; Gilboa & Schmeidler 1989; Mukerji & Tallon 2001; Wald 1950) for uncertainty. For these theories simple techniques to measure their primitives empirically have not yet been discovered.

4. THE EXPERIMENT: METHOD

Participants. N=78 undergraduate students participated from a wide range of disciplines at the University of Amsterdam.

Procedure. Participants were seated in front of personal computers in seven different sessions with approximately 11 participants per session. Participants first received experimental instructions (see Appendix A), after which the experimental questions followed.

Stimuli; general. Participants were asked two practice choice questions to familiarize them with the experimental procedures. The choice questions involved simple choices between two prospects named prospect L (left) and prospect R (right). Both prospects yielded prizes depending on the outcome of a roll with two ten-sided dice, generating probabilities j/100.[1] Prospects were framed as in Figure 4. Participants were asked to indicate their choice by simply clicking on the button representing their preferred prospect and were encouraged to answer the choice questions at their own pace. In order to avoid potential confounding effects resulting from connotations with words such as lottery or gamble, we used the more neutral term prospect in the instructions. The position of each of the two prospects was counterbalanced between participants in order to avoid a potential confounding effect that might result from individual preference for a particular position of a prospect.

FIGURE 4. The framing of the prospect pairs

PROSPECT L		
roll	probability	prize
1 till p	p %	x_{i-1} euro
p+1 till 100	(100 – p)%	Y euro

PROSPECT R		
roll	probability	prize
1 till p	p %	x_i euro
p+1 till 100	(100 – p)%	y euro

Stimuli of the part measuring utility. In the “outcome part” of the experiment, we set $x_0 = 60$ and obtained values x_1 and x_2 to generate indifferences $(0.25{:}x_1, 30) \sim (0.25{:}60, 40)$ and $(0.25{:}x_2, 30) \sim (0.25{:}x_1, 40)$. Then x_1 is the utility midpoint between x_0 and x_2 (Section 2). Because all further measurements in the experiment depended on the values x_1 and x_2, these values were elicited twice and the average of the two values obtained was used as input in the rest of the experiment, so as to reduce noise. To obtain indifferences we used a bisection choice method. This method, while time-consuming, has been found to give more consistent results (Bostic, Herrnstein & Luce 1990) than direct matching.

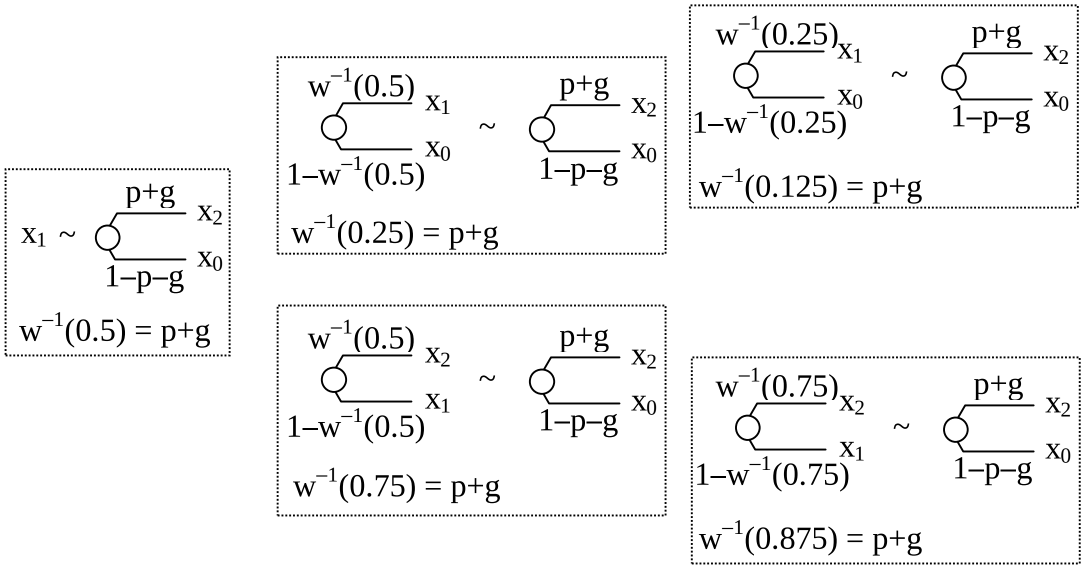
Our bisection method is similar to the method used by Abdellaoui (2000), and was as follows. To obtain x_1 to generate the indifference $(0.25{:}\ x_1, 30) \sim (0.25{:}\ x_0, 40)$, we iteratively narrowed down so-called indifference intervals (containing x_1), as follows. Based on extensive pilots, we hypothesized that x_1 would not exceed $x_0 + 96$ and took $[x_0, x_0+96]$ as the first indifference interval, denoted $[\ell^1,u^1]$. To construct

[1] One 10-sided die was numbered from 0 till 9 while the other 10-sided die was numbered from 00 till 90. Because we informed participants that the roll 0-00 would be coded as 100, the sum of a roll with both 10-sided dice resulted in a random number ranging from 1 till 100.

the j+1th indifference interval from the jth indifference interval $[\ell^j, u^j]$, we elicited whether the midpoint $\frac{\ell^j+u^j}{2}$ of $[\ell^j, u^j]$ is larger or smaller than x_1. To do so, we observed the choice between $(0.25: \frac{\ell^j+u^j}{2}, 30)$ and $(0.25: x_0, 40)$. A left choice means that the midpoint is larger than x_1, so that x_1 is contained in $[\ell^j, \frac{\ell^j+u^j}{2}]$, which was then defined as the j+1th indifference interval $[\ell^{j+1}, u^{j+1}]$. A right choice means that the midpoint is smaller than x_1, so that x_1 is contained in $[\frac{\ell^j+u^j}{2}, u^j]$, which was then defined as the j+1th indifference interval $[\ell^{j+1}, u^{j+1}]$. We did five iteration steps like this, ending up with $[\ell^6, u^6]$ (of length $96 \times 2^{-5} = 3$), and took its midpoint as the elicited indifference value x_1. We similarly elicited x_2 (substitute x_2 for x_1 and x_1 for x_0 above).

Stimuli of the part measuring probability weighting. In the "probability part" of the experiment, we employed our measurement technique to obtain five probabilities that are equally spaced in terms of subjective probability weights. We will denote these five probabilities by $w^{-1}(.125)$, $w^{-1}(.25)$, $w^{-1}(.5)$, $w^{-1}(.75)$ and $w^{-1}(.875)$, where $w^{-1}(s)$ is the probability corresponding to a subjective probability weight of s. Again, we derived indifferences from binary choices and framed the prospects as in Figure 4. All left prospects used in the experiment are special cases of Prospect L in Figure 1 with one probability 0, so that only two branches remain.

FIGURE 5. The elicited indifferences



As shown in Section 3, our elicitation technique concerns indifference between prospect L = (p:x_2, d:x_1, r: x_0) and prospect R = (g+p: x_2, r+b: x_0), implying that probability $g+p$ is the weight midpoint between probability p and probability $d+p$. For example, to obtain $w^{-1}(.5)$, the weight midpoint between 0 and 1, we take p=0 and d=1, so that L is the degenerate prospect yielding x_1 with certainty. Figure 5 lists the elicited indifferences used to obtain the probabilities $w^{-1}(.125)$, $w^{-1}(.25)$, $w^{-1}(.5)$, $w^{-1}(.75)$, and $w^{-1}(.875)$.

In general, to find g to generate an indifference (p:x_2, d:x_1, r:x_0) ~ (g+p:x_2, x_0) as in Figure 1, we used a bisection method as in the outcome part of the experiment. We iteratively narrowed down so-called indifference intervals containing $g+p$, as follows. The first indifference interval $[\ell^1, u^1]$ was [p, d+p], i.e. the interval of which the weighting-midpoint was to be found.[2] By stochastic dominance, it contains $g+p$ indeed. Each participant was first asked to make two practice choices between a particular prospect L and the prospect R = (g^++p:x_2, x_0) (R = (g^-+p:x_2, x_0)), where probability g^++p (g^-+p) was set equal to the upper (lower) limit of the range of the first indifference interval of probability $g+p$ minus (plus) 1/100. Then the iterative process started.

[2] The first indifference interval is, thus, [0,1] for $w^{-1}(0.5)$, [0, $w^{-1}(0.5)$] for $w^{-1}(0.25)$, [$w^{-1}(0.5)$, 1] for $w^{-1}(0.75)$, [0, $w^{-1}(0.25)$] for $w^{-1}(0.125)$, and [$w^{-1}(0.75)$, 1] for $w^{-1}(0.875)$.

To construct the j+1[th] indifference interval $[\ell^{j+1},u^{j+1}]$ from the j[th] indifference interval $[\ell^{j},u^{j}]$, we elicited whether the midpoint of $[\ell^{j},u^{j}]$ is larger or smaller than p+g. To do so, we observed the choice between $(p{:}x_2, d{:}x_1, r{:}x_0)$ and $(\frac{\ell^j+u^j}{2}{:}x_2, x_0)$. A right choice means that the midpoint is larger than g+p, so that g+p is contained in $[\ell^j,\frac{\ell^j+u^j}{2}]$, which was then defined as the j+1[th] indifference interval $[\ell^{j+1},u^{j+1}]$. A left choice means that the midpoint is smaller than g+p, so that g+p is contained in $[\frac{\ell^j+u^j}{2},u^j]$, which was then defined as the j+1[th] indifference interval $[\ell^{j+1},u^{j+1}]$. We did five iteration steps like this, ending up with $[\ell^6,u^6]$, and took its midpoint as the elicited indifference probability p+g.

Because prospects yielded prizes depending on the result of a roll with two ten-sided dice, we only allowed values j/100 for probabilities. When a particular midpoint probability was not a value j/100, the computer took the closest value j/100 on the left of this value if the value was lower than half and on the right of this value if the value was higher than half. The order of elicitation was varied between participants to prevent potential order effects. For some participants the order of elicitation was $w^{-1}(.5)$, $w^{-1}(.25)$, $w^{-1}(.75)$, $w^{-1}(.125)$, $w^{-1}(.875)$, whereas for other participants the order of elicitation was $w^{-1}(.5)$, $w^{-1}(.75)$, $w^{-1}(.25)$, $w^{-1}(.875)$, $w^{-1}(.125)$.

As an illustration, Figure 6 replicates the bisection procedure followed to obtain the probability corresponding to the weight of 0.5. The particular pattern of answers depicted there, preferring the right prospect twice and the left prospect three times, was exhibited by 6 of our participants. After the fifth iteration step, the midpoint of the last indifference interval was taken as the final indifference probability. Thus, individual indifference between the certain prospect (x_1) and the prospect $(.615{:}x_2, x_0)$ was inferred from the choices made by the 6 participants whose choices are replicated in Figure 6.

FIGURE 6. The bisection method for measuring $w^{-1}(0.5)$

Choice Question	Indifference interval	Prospect L	Prospect R	Choice
1	$[\ell^1,u^1] = [0, 1]$	x_1	0.50 x_2; 0.50 x_0	L
2	$[\ell^2,u^2] = [0.50, 1]$	x_1	0.75 x_2; 0.25 x_0	R
3	$[\ell^3,u^3] = [0.50, 0.75]$	x_1	0.63 x_2; 0.37 x_0	R
4	$[\ell^4,u^4] = [0.50, 0.63]$	x_1	0.57 x_2; 0.43 x_0	L
5	$[\ell^5,u^5] = [0.57, 0.63]$	x_1	0.60 x_2; 0.40 x_0	L

CONCLUSION: $[\ell^6,u^6] = [0.60, 0.63]$; $x_1 \sim$ 0.615 x_2; 0.385 x_0; $w^{-1}(0.5) = 0.615$.

Motivating participants. We used performance-based real incentives to motivate the participants, based on the random-lottery incentive system, the nowadays almost exclusively used real-incentive system for individual choice experiments (Holt & Laury 2002), as follows. For each session there were as many envelopes as participants, with one envelop containing a blue card and all others a white card. Each participant could choose an envelope, after which the participant who got the blue card could play for real. For this participant, one choice question was again selected randomly and the chosen prospect in that choice question was played out for real, with the participant paid according to the prospect chosen and the outcome that resulted from playing out this prospect. All other participants in a particular session, who had chosen a white card, received a fixed payment of €5. The possible monetary outcomes of the prospects used during the experiment ranged from €30 to approximately €250. All payments were done privately and immediately at the end of the experiment. The average payment under real play was €77.57, so that the total reward per participant was approximately $10/11 \times 5 + 1/11 \times 77.57$ = €11.60, while it took participants about 20 minutes to complete the experiment.

Further Stimuli. Our questions were chained, and it is well-known that chaining can give incentives for not truly answering questions (Harrison 1986). To check out if participants had been aware of this possibility, we asked two *strategy-check questions*: "Was there any special reason for you to specially choose left more often, or specially choose right more often?" and "Can you state briefly which method you used to determine your choice?" These questions were asked in a questionnaire at the end of the experiment, with further questions about age, study, and gender.

5. THE EXPERIMENT: RESULTS

14 participants were excluded from the analysis because they gave erratic or heuristic answers such as always choosing the left option or always choosing the right option. The practice choices of this experiment also served to detect such erratic and heuristic answers. These participants apparently did not understand the choices or did not seriously think about them. Including these participants in the analysis would not alter the results presented hereafter.

In the the strategy-check questions, no participant revealed awareness of the chained nature of the questions, or an attempt to strategically exploit this chaining. 25 subjects indicated a combination of (expected or maximal) value and safety, 5 went merely by expected value, 4 merely by highest value, and various other reasons were given.

5.1. The Utility Function

The first measurement of outcome x_1 (x_2) did not differ significantly from the second measurement (Wilcoxon signed-rank tests, $z = 1.23$, $p = 0.2$ and $z = -1.48$, $p = 0.14$). We, therefore, take averages of the two measurements in the following analyses (as we did for the stimuli during the experiment).

The median values of x_1 and x_2 are 92.25 and 123, respectively, which, together with $x_0 = 60$, suggests linear utility. The deviation from linearity is not statistically significant (Wilcoxon signed-rank test, $z = 0.887$, $p = 0.3751$), in agreement with the common hypothesis that utility is approximately linear for moderate amounts of money.

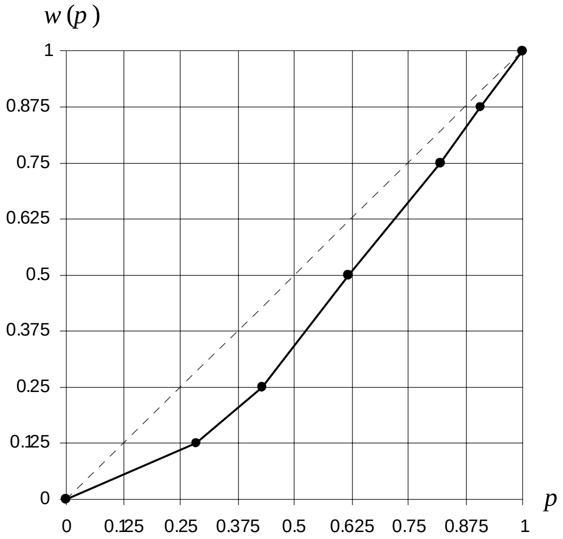
5.2. The Probability Weighting Function

5.2.1. Non-Parametric Analysis

There was no significant effect of the order of elicitation of decision weights and we, hence, pooled the data. Figure 7 displays both the weighting function based on median values and the corresponding summary statistics. Overall we find a convex (pessimistic) pattern; the median values of $w^{-1}(.125)$, $w^{-1}(.25)$, $w^{-1}(.5)$, $w^{-1}(.75)$ and $w^{-1}(.875)$ were .285, .430, .608, .793, and .910, respectively. Table 1 shows that participants did not process probabilities linearly, but mostly underweighted probabilities. The differences between the obtained probabilities for the different probability weights and the probabilities corresponding to a linear probability weighting function are all highly significant, except for $w^{-1}(.875)$.

We classified participants on the basis of the shape of their probability weighting function by calculating *slope differences*, i.e. the change in the average slope of the probability weighting function between adjacent probability intervals. There are five slope differences for each participant. We used a categorization similar to Bleichrodt & Pinto (2000), adapted to our context: The probability weighting function of a participant was classified as convex (concave; linear) if at least three slope differences were positive (negative; zero); see Table 2.

FIGURE 7 – The Median Probability Weighting Function



Summary Statistics

$w^{-1}(p)$	Mean	Median	Standard Deviation
0.125	0.330	0.285	0.228
0.250	0.441	0.430	0.223
0.500	0.608	0.620	0.193
0.750	0.793	0.820	0.150
0.875	0.872	0.910	0.132

TABLE 1 – Counts of $w^{-1}(p) - p > 0$ and $w^{-1}(p) - p < 0$

$w^{-1}(p) - p$	> 0	< 0
p = 0.125	49**	15
p = 0.250	48**	16
p = 0.500	44**	20
p = 0.750	44**	18
p = 0.875	41	23

**denotes significance at the 1% level using a two-tailed Wilcoxon signed-rank test.

TABLE 2 – Classification of Participants

Shape	% of Participants
Concave	25 %
Convex	62.5 %
Linear	0 %
Unclassified	12.5%

This classification again suggests that the weighting function was predominantly convex.

5.2.2. Parametric Analysis

Several functional forms of the probability weighting function have been proposed in the literature. The most popular one-parameter specifications are the functionals proposed by Tversky & Kahneman (1992) and Prelec (1998). The most popular two-parameter functional forms are the ones proposed by Goldstein & Einhorn (1987) and Prelec (1998). The power family has lost popularity. The second column of Table 3 lists the parametric specifications proposed by the aforementioned authors. The following results will illustrate clearly that patterns found can be driven more by the parametric family chosen than by the actual data.

TABLE 3. Parameter Estimates

Study	w(p)	Median estimate	Individual distance	Distance from median
Hey & Orme (1994)	p^{γ}	γ = 1.5041 (0.28)	0.0515	0.0019
Tversky & Kahneman (1992)	$\frac{p^{\gamma}}{(p^{\gamma}+(1-p)^{\gamma})^{1/\gamma}}$	γ = 2.1626 (1.24)	0.0556	0.0037
Goldstein & Einhorn (1987)	$\frac{\delta p^{\gamma}}{\delta p^{\gamma}+(1-p)^{\gamma}}$	γ = 1.2379 (0.10) δ = 0.4120 (0.12)	0.0153	0.0031
Prelec(1) (1998)	$e^{-(-\ln p)^{\alpha}}$	α = 0.5206 (0.26)	0.1174	0.0752
Prelec(2) (1998)	$e^{-\beta(-\ln p)^{\alpha}}$	α = 1.0540 (0.06) β = 1.7630 (1.64)	0.0161	0.0022

Standard errors are in parentheses.

For each participant and each parametric specification, we estimated the optimal parameter values by minimizing sums of squared distances:

$$\sum_{i=1}^{5}(w_i - \hat{w}_i)^2. \tag{5.1}$$

Here w_i is the i-th element of the sequence of the probability weights for which the probabilities were elicited and $\hat{w}_i$ is the i-th element of the estimated sequence of probability weights under the various parametric specifications. For example, for $w_1 = 0.125$, we define p_1 as the probability elicited in the experiment as $w^{-1}(0.125)$ for this subject, and then $\hat{w}_1$ is the value that the probability weighting function of the parametric family assigns to p_1. Its distance from 0.125 indicates how far the family is from the data.

To avoid convergence to local minima we used a wide variety of starting points. The medians of the individual estimators as well as the standard errors are reported in the third column of Table 3 and the corresponding weighting functions are plotted in Figure 8. Table 3 also gives the average squared distances if for each individual the optimal parameters are determined, as well as the squared distances of the optimal fits to the median data.

FIGURE 8. Parametric fittings of median data

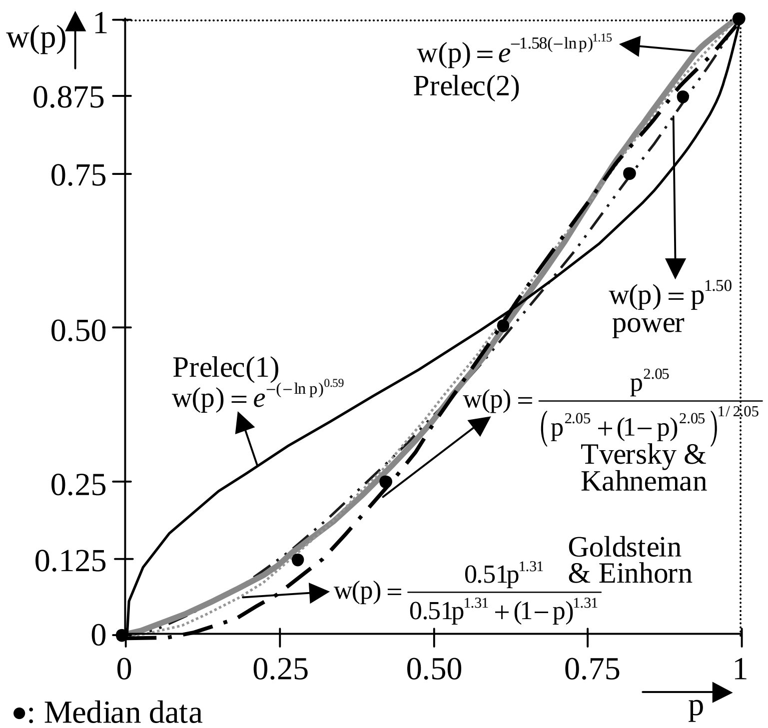


Figure 8 displays a remarkable variety of patterns, which should caution against the unqualified use of parametric fitting. The one-parameter family advocated by Prelec (1998), denoted Prelec(1), displays the most commonly found shape, the inverse-S shape. It is, however, clear that the function is completely off, and its distances from the data in Table 3 are extreme. This family necessarily imposes the inverse-S shape and intersects the diagonal at (1/e,1/e), which explains part of its bad performance. The one-parameter family by Tversky & Kahneman (1992) and the two-parameter families by Prelec (1998) and Goldstein & Einhorn (1987) show a similar pattern: mostly convex but, being oriented to inverse-S like phenomena, they go the other way and to some extent display the opposite pattern, being a slight S shape. Finally, the power family yields a convex function.

The Individual Sum Squared Residuals suggest that of the one-parameter families, the one of Tversky & Kahneman performs considerably better than Prelec's, but the power family is best. The two two-parameter families perform very similarly, and have a smaller distance than the one-parameter families which is no surprise given their larger number of free parameters. The optimal fits for the median data give similar results with one exception: the power-family now yields the best fit, even better than the families with an extra parameter do.

6. DISCUSSION

We used the tradeoff measurement technique of Wakker & Deneffe (1996) to obtain utility midpoints derived endogenously from preference, as suggested by Köbberling & Wakker (2003, p. 408). Abdellaoui, Bleichrodt, & Paraschiv (2005) similarly used this suggestion. They next obtained a probability q with w(q) = 0.5 through what amounts to a degenerate version of Figure 1 with r=1 and p=0. Finally, they used this probability to efficiently measure utility midpoints in general. Their approach, like ours, can be interpreted as a special case of Blavatskyy's (2006) general procedure.

Vind (1991, p. 134; 2003, Section IV.2, above Theorem IV.2.1) proposed an alternative method for obtaining endogenous utility midpoints under expected utility and, more generally, under state-dependent expected utility (from which he derived a so-called mean groupoid operation). He showed that y is the utility midpoint between x and z if the following indifferences hold:

$$x \sim (p{:}x_1,x_2),\ z \sim (p{:}z_1,z_2), \text{ and } (p{:}x_1,z_2) \sim (p{:}z_1,x_2) \sim y. \quad (6.1)$$

His method holds under prospect theory if we add the requirement that $x_1 > x_2$, $x_1 > z_2$, $z_1 > z_2$, and $z_1 > x_2$.

Ghirardato et al. (Definition 4) proposed another method to derive utility midpoints endogenously from preferences. They showed that β is the utility midpoint between α and γ under prospect theory if the following indifferences hold:

$$(p{:}\alpha, \gamma) \sim (p{:}x,y),\ x \sim (p{:}\alpha,\beta), \text{ and } y \sim (p{:}\beta,\gamma) \quad (6.2)$$

with $\alpha > \beta > \gamma$. We chose the tradeoff method because it requires the measurement of fewer indifferences and is easier to implement than the two methods just discussed.

The answers x_1, x_2, and $w^{-1}(p)$ that were elicited from participants returned as inputs in later questions (chaining), and bisection also involves chaining. It is well known that participants can exploit chaining by not answering truthfully at particular questions so as to improve stimuli in future questions (Harrison 1986). Such a distortion is unlikely to have arisen in our experiment. It is difficult for participants to understand that their anwer to one question will influence future stimuli. For example, we did not directly ask for the indifference values used in future questions, but derived indifference values indirectly from the choices made, so that participants had not seen these values before and in this way could not recognize them. In addition, to exploit chaining, not only the presence of chaining must be understood, but also the way in which future questions depend on current answers. Our strategy-check question was designed to detect strategic exploitation of chaining if present, with the wording of the question chosen so as not to influence participants, but none of the participant revealed such exploitation. In our instructions for the random-lottery incentive system we chose our wording very carefully so as not to write anything untrue (see end of Instructions in Appendix A).

To avoid lying to participants, we carefully formulated our instructions (end of Appendix A) as regards chaining. Approaches where participants are not told lies but do not receive complete information about the whole setup of the experiment, are not uncommon in experimental economics. For instance, in Holt & Laury (2002), participants were told that they would be paid for questions in a first part of the experiment (the “small-payment treatment”), which was true in a literal sense. They were, however, not told that afterwards, in case they would continue to the second part of the experiment with a large-payment treatment, an event that was obvious to happen, their payments of the first part would be taken away. Thus, it was obvious beforehand that participants would end up not being paid for the first part of the experiment, but they were not informed about that.Empirical studies have found that weighting functions are mostly convex or inverse-S shaped, with the latter shape prevailing. In our study, the former shape was prevailing. This finding suggests once more that probability weighting is a volatile phenomenon, with results depending on framing and ways of measurement, and no phenomena holding in great generality.

7. CONCLUSION

We have introduced a new way to measure probability weighting functions of rank-dependent utility (Gilboa 1987; Quiggin 1981; Schmeidler 1989) and prospect theory (Tversky & Kahneman 1992), both for risk and for uncertainty, and implemented it in an experiment for risk. Our results suggest that the weighting function may be more volatile and dependent on framing and measurement techniques than thought before. In particular, the result of parametric fitting can depend much on the parametric family chosen. The experiment demonstrated that our method is feasible and more efficient than methods used before.

APPENDIX A. EXPERIMENTAL INSTRUCTIONS

[Instructions are translated from Dutch]

Welcome at this experiment. If you have any question while reading these instructions, please raise your hand. The experimenter will then come to your table to answer your question. This experiment takes about half an hour. We ask you to make a number of decisions during this experiment. Each time, you choose between two so-called “prospects.” Both prospects yield certain prizes depending on the roll of the two ten-sided dice similar to the ones that are lying on your table right now.

As you can see, one 10-sided die has the values 0, 1, 2, 3, 4, 5, 6, 7, 8 and 9 and the other 10-sided die has the values 00, 10, 20, 30, 40, 50, 60, 70, 80 and 90. If we code the sum of the roll “a 0 and a 00” as 100, the sum of a roll with both 10-sided dice thus yields a random number from 1 up until 100.

The prospects from which you have to choose named Prospect L (left) and Prospect R (right) will be presented in the following way:

PROSPECT L		
roll	probability	prize
1 till 40	40%	100 euro
41 till 100	60%	50 euro

PROSPECT R		
roll	probability	prize
1 till 20	20%	150 euro
21 till 100	80%	20 euro

In this case, Prospect L yields a prize of 100 Euro if the sum of the roll with both 10-sided dice is 1 up until 40 and if the sum of a roll is 41 up until 100, Prospect L yields a prize of 50 Euro, as you can see. Similarly, Prospect R yields a prize of 150 Euro if the sum of a roll with both 10-sided dice is 1 up until 20 and otherwise Prospect R yields a prize of 20 Euro, in this case.

Both the prizes as well as the probabilities of yielding certain prizes can vary across decisions. We ask you to choose between Prospect L and Prospect R each time, by clicking the corresponding button with the mouse.

For your participation in this experiment, you receive 5 Euro anyhow. Thereby, one participant of this experiment will be selected at random at the end of this experiment. For this purpose, each participant randomly picks a sealed envelope containing either a white or a blue card at the end of the experiment. Participants selecting an envelope containing a white card receive 5 Euro for their participation. From the participant with an envelope containing a blue card one decision will be selected at random by rolling both 10-sided dice. Thereafter, the prize of the chosen prospect in that decision will be determined by rolling the two 10-sided dice again. This prize, always larger than 5 Euro, will be paid out to the particular participant.

There are no right or wrong answers during this experiment. It exclusively concerns your own preferences. In those we are interested. At every decision it is best for you to choose the prospect that you want most. Surely, if you select the envelope containing the blue card at the end of the experiment, that decision can be selected at the end of the experiment. Then, the chosen prospect will be played out. Off course, you would like that prospect to be the prospect you want most. If you have no further questions you can now start with the experiment by clicking on the "Continue" button below.

REFERENCES

Abdellaoui, Mohammed (2000), "Parameter-Free Elicitation of Utilities and Probability Weighting Functions," *Management Science* 46, 1497–1512.

Abdellaoui, Mohammed, Han Bleichrodt, & Corina Paraschiv (2004), "Measuring Loss Aversion under Prospect Theory: A Parameter-Free Approach," iMTA/iBMG, Erasmus University, Rotterdam, The Netherlands.

Bell, David E. (1982), "Regret in Decision Making under Uncertainty," *Operations Research* 30, 961–981.

Blavatskyy, Pavlo R. (2006), "Error Propagation in the Elicitation of Utility and Probability Weighting Functions," *Theory and Decision* 60, 315–334.

Bleichrodt, Han & José Luis Pinto (2000), "A Parameter-Free Elicitation of the Probability Weighting Function in Medical Decision Analysis," *Management Science* 46, 1485–1496.

Bostic, Raphael, Richard J. Herrnstein, & R. Duncan Luce (1990), "The Effect on the Preference-Reversal Phenomenon of Using Choice Indifferences," *Journal of Economic Behavior and Organization* 13, 193–212.

Chateauneuf, Alain (1991), “On the Use of Capacities in Modeling Uncertainty Aversion and Risk Aversion,” *Journal of Mathematical Economics* 20, 343–369.

Chew, Soo Hong (1983), “A Generalization of the Quasilinear Mean with Applications to the Measurement of Income Inequality and Decision Theory Resolving the Allais Paradox,” *Econometrica* 51, 1065–1092.

Chew, Soo Hong & Guofu Tan (2005), "The Market for Sweepstakes," *Review of Economic Studies* 72, 1009–1029.

Dekel, Eddie (1986), “An Axiomatic Characterization of Preferences under Uncertainty: Weakening the Independence Axiom,” *Journal of Economic Theory* 40, 304–318.

Donkers, A.C.D., Bertrand Melenberg, & Arthur H.O. van Soest (2001), “Estimating Risk Attitudes Using Lotteries; A Large Sample Approach,” *Journal of Risk and Uncertainty* 22, 165–195.

Dybvig, Philip & Heracles M. Polemarchakis (1981), “Recovering Cardinal Utility,” *Review of Economic Studies* 48, 159–166.

Etchart, Nathalie (2004), “Is Probability Weighting Sensitive to the Magnitude of Consequences? An Experimental Investigation on Losses,” *Journal of Risk and Uncertainty* 28, 217–235.

Ghirardato, Paolo, Fabio Maccheroni, Massimo Marinacci, & Marciano Siniscalchi (2003), “A Subjective Spin on Roulette Wheels,” *Econometrica* 71, 1897–1908.

Gilboa, Itzhak (1987), “Expected Utility with Purely Subjective Non-Additive Probabilities,” *Journal of Mathematical Economics* 16, 65–88.

Gilboa, Itzhak & David Schmeidler (1989), “Maxmin Expected Utility with a Non-Unique Prior,” *Journal of Mathematical Economics* 18, 141–153.

Gold, Marthe R., Joanna E. Siegel, Louise B. Russell, & Milton C. Weinstein (1996), “*Cost-Effectiveness in Health and Medicine.*” Oxford University Press, New York.

Goldstein, William M. & Hillel J. Einhorn (1987), “Expression Theory and the Preference Reversal Phenomena,” *Psychological Review* 94, 236–254.

Gonzalez, Richard & George Wu (1999), “On the Shape of the Probability Weighting Function,” *Cognitive Psychology* 38, 129–166.

Gul, Faruk (1991), “A Theory of Disappointment Aversion,” *Econometrica* 59, 667–686.

Harrison, Glenn W. (1986), “An Experimental Test for Risk Aversion,” *Economic Letters* 21, 7–11.

Hey, John D. & Chris Orme (1994), “Investigating Generalizations of Expected Utility Theory Using Experimental Data,” *Econometrica* 62, 1291–1326.

Holt, Charles A. & Susan K. Laury (2002), “Risk Aversion and Incentive Effects,” *American Economic Review* 92, 1644–1655.

Jullien, Bruno & Bernard Salanié (2000), “Estimating Preferences under Risk: The Case of Racetrack Bettors,” *Journal of Political Economy* 108, 503–530.

Keeney, Ralph L. & Howard Raiffa (1976), “*Decisions with Multiple Objectives.*” Wiley, New York (Second edition 1993, Cambridge University Press, Cambridge, UK).

Köbberling, Veronika & Peter P. Wakker (2003), “Preference Foundations for Nonexpected Utility: A Generalized and Simplified Technique,” *Mathematics of Operations Research* 28, 395–423.

Loomes, Graham & Robert Sugden (1982), “Regret Theory: An Alternative Theory of Rational Choice under Uncertainty,” *Economic Journal* 92, 805–824.

Mukerji, Sujoy & Jean-Marc Tallon (2001), “Ambiguity Aversion and Incompleteness of Financial Markets,” *Review of Economic Studies* 68, 883–904.

Prelec, Drazen (1998), “The Probability Weighting Function,” *Econometrica* 66, 497–527.

Quiggin, John (1981), “Risk Perception and Risk Aversion among Australian Farmers,” *Australian Journal of Agricultural Economics* 25, 160–169.

Rabin, Matthew (2000), “Risk Aversion and Expected-utility Theory: A Calibration Theorem,” *Econometrica* 68, 1281–1292.

Schmeidler, David (1989), “Subjective Probability and Expected Utility without Additivity,” *Econometrica* 57, 571–587.

Starmer, Chris (1992), “Testing New Theories of Choice under Uncertainty Using the Common Consequence Effect,” *Review of Economic Studies* 59, 813–830.

Tversky, Amos & Craig R. Fox (1995), “Weighing Risk and Uncertainty,” *Psychological Review* 102, 269–283.

Tversky, Amos & Daniel Kahneman (1981), “The Framing of Decisions and the Psychology of Choice,” *Science* 211, 453–458.

Tversky, Amos & Daniel Kahneman (1992), “Advances in Prospect Theory: Cumulative Representation of Uncertainty,” *Journal of Risk and Uncertainty* 5, 297–323.

Vind, Karl (1991), “Independent Preferences,” *Journal of Mathematical Economics* 20, 119–135.

Vind, Karl (2003), “*Independence, Additivity, Uncertainty.*” With contributions by B. Grodal. Springer, Berlin. First version 1969.

Wakker, Peter P. & Daniel Deneffe (1996), “Eliciting von Neumann-Morgenstern Utilities when Probabilities Are Distorted or Unknown,” *Management Science* 42, 1131–1150.

Wald, Abraham (1950), *“Statistical Decision Functions.”* Wiley, New York.

Young, H. Peyton (1990), “Progressive Taxation and Equal Sacrifice,” *American Economic Review* 80, 253–266.