

1                   A Midpoint Technique for Easily  
2                   Measuring Prospect Theory's Probability  
3                   Weighting\*

4                   GIJS VAN DE KUILEN<sup>a</sup>, PETER P. WAKKER<sup>b</sup>, & LIANG ZOU<sup>a</sup>

5                                   August, 2006

6                   *a: Dept. of Economics, University of Amsterdam, Roetersstraat 11, Amsterdam, 1018*

7                                   *WB, The Netherlands;*

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

11  
12   ABSTRACT. Prospect theory can better describe risky choices than classical expected  
13   utility because it makes the plausible assumption that risk aversion is driven as much  
14   by the way people feel about probabilities (probability weighting) as about outcomes  
15   (utility). This leads to better predictions but, as a price to pay, probability weighting  
16   is more difficult to measure than utility, which may explain why many economists  
17   today continue to use expected utility. This paper mitigates the drawback mentioned  
18   by introducing a new method for measuring probability weighting that is simpler and  
19   more efficient than methods used before. The new method is implemented in an  
20   experiment. Most participants exhibited a convex weighting function, implying  
21   pessimism and enhancing risk aversion. This finding supports recent claims that  
22   utility is less concave than was traditionally thought in studies that ascribed all risk  
23   aversion to concave utility.

24  
25  
26   JEL-Classification: D81, C60, C91

27   KEYWORDS: Prospect Theory; Probability Weighting; Pessimism

---

\* Han Bleichrodt made useful comments.

28

## 1. INTRODUCTION

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

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

53 Our method is based on a technique for measuring midpoints of decision  
54 weights. For simplicity of presentation, we will present our new technique in an  
55 experiment for decision under risk (known probabilities). As we will demonstrate in  
56 the theoretical analysis, our measurement method can equally well be used to measure  
57 nonadditive weighting functions for uncertainty with unknown probabilities (Gilboa  
58 1987; Schmeidler 1989; Tversky & Kahneman 1992). Hence, it can also be used to  
59 examine ambiguity attitudes.

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

71 We use our method in an experiment with  $N=78$  participants. In the literature,  
 72 the most common finding is a combination of inverse-S shaped and convex-shaped  
 73 curves, where the inverse-S shape has been found to be prevailing in most studies  
 74 (Abdellaoui 2000; Bleichrodt & Pinto 2000; Gonzalez & Wu 1999; Tversky & Fox  
 75 1995; Tversky & Kahneman 1992). Convexity was assumed for instance in Starmer  
 76 (1992, p. 818 and pp. 819-820, expressed as concavity of the dual of  $w$ ), and convex  
 77 shapes were found to be prevailing in Jullien and Salanié (2000). In the present study  
 78 we find both shapes too, with convex shapes prevailing.

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

## 85 2. PROSPECT THEORY AND EARLIER MEASUREMENT 86 TECHNIQUES

87 Outcomes are monetary, with  $\mathbb{R}^+$  the outcome set. We do not consider  
 88 negative outcomes (losses). A *prospect* is a probability distribution over outcomes  
 89 taking only finitely many outcomes;  $(p_1: x_1, \dots, p_n: x_n)$  denotes a prospect yielding

90 outcome  $x_i$  with probability  $p_i$ ,  $i = 1, \dots, n$ , where probabilities are nonnegative and  
 91 sum to one. For two outcomes we often suppress the second probability, and write  
 92  $(p:x, y)$  for  $(p:x, 1-p:y)$ .  $U: \mathbb{R}^+ \rightarrow \mathbb{R}$  denotes the utility function. It is continuous and  
 93 strictly increasing. Probabilities are transformed into probability weights through a  
 94 strictly increasing and continuous *probability weighting function*  $w: [0,1] \rightarrow [0,1]$ ,  
 95 with  $w(0) = 0$  and  $w(1) = 1$ .

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

$$104 \quad \pi_i = w(p_i + \dots + p_1) - w(p_{i-1} + \dots + p_1), \quad (2.1)$$

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

$$108 \quad \sum_{i=1}^n \pi_i U(x_i). \quad (2.2)$$

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

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

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

127 Gonzalez & Wu (1999) did not commit to a parametric family but still used  
 128 fitting techniques that minimize squared distances, based on a complex numerical  
 129 system that requires much data per participant. In return, their results are very  
 130 reliable. Abdellaoui (2000) and Bleichrodt & Pinto (2000) provided two more  
 131 tractable methods for estimating probability weighting functions nonparametrically.  
 132 Details are given below. As with all other measurements used so far, these methods  
 133 need a detailed measurement of utility before probability weighting can be measured.  
 134 The present paper introduces a new and simpler technique that, for each pair of  
 135 probabilities, can easily infer the midpoint probability in terms of probability weights.  
 136 It then becomes easy to measure the probability weighting function to any degree of  
 137 precision. Starting from  $w(0) = 0$  and  $w(1) = 1$ , we can measure the  $p$  with  $w(p) =$   
 138  $1/2$ , i.e.  $w^{-1}(1/2)$ . Then, with  $w(0) = 0$  and  $w^{-1}(1/2)$  available, we can measure  
 139  $w^{-1}(1/4)$ , and, similarly,  $w^{-1}(3/4)$ ,  $w^{-1}(1/8)$ ; etc. We can continue measuring  
 140 midpoints until we have estimated the curve of  $w$  as accurately as we want.

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

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

$$150 \quad (p: x_{i+1}, y) \sim (p: x_i, Y), \text{ for } x_{i+1} > x_i > Y > y, \quad i = 0, \dots, n. \quad (2.3)$$

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

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

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

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

160  $x_i \sim (p_i: x_6, x_0)$  for all  $i$ .

161 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

162  $w(p_i) = i/6$  for all  $i$ .

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

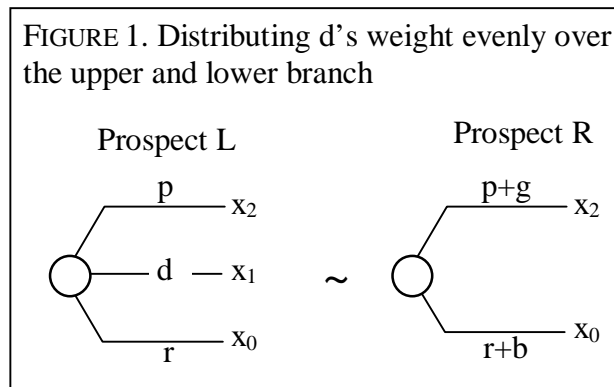
166 These methods make extensive use of the outcome domain and need to carry  
 167 out a utility measurement at least as detailed as the probability weighting  
 168 measurement that is desired. For example, to measure the probability with weight  
 169  $1/32$ , we need to elicit 32 values  $x_i$ , or rely on parametric interpolations. Our method  
 170 will avoid these complications. Like the methods described, it uses information  
 171 obtained from utility measurements to elicit decision weights, but it does so in a  
 172 different and more efficient manner. Blavatsky (2006) described the general  
 173 procedure to start with measurements in one dimension, then use this to obtain  
 174 measurements in the other dimension, possibly using these again to obtain  
 175 measurements in the first dimension, and so forth. He examined general efficiency  
 176 principles of such general procedures.

177

### 178 3. A NEW MIDPOINT TECHNIQUE

179 Our method of measuring midpoints of a weighting function starts with  
 180 measuring a midpoint of utility. To this end, we measure  $x_0, x_1, x_2$  as in Eq. 2.3, after

181 which  $x_1$  results as the midpoint between  $x_0$  and  $x_2$  in utility units. These  $x$ -values  
 182 will be used throughout what follows. Alternative methods for endogenously deriving  
 183 utility midpoints from preferences by Ghirardato et al. (2003) and Vind (2003) will be  
 184 discussed in Section 6. To elicit weighting functions, we elicit indifferences between  
 185 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  
 186 residual probability  $1-p-d$ . Here  $d$  is the probability mass of  $x_1$  to be divided over the  
 187 other outcomes.  $g$  is the probability mass taken from  $d$  and moved to the good  
 188 outcome  $x_2$ , and the remainder  $b = d-g$  is moved to the bad outcome  $x_0$ .



197

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

208

209 **Theorem 1.** The indifference in Figure 1 implies that

210 
$$w(g + p) = \frac{w(p) + w(d+p)}{2}$$

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

212

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

$$215 \quad w(p)(U(x_2)-U(x_0)) + (w(d+p) - w(p))(U(x_1) - U(x_0)) = w(g+p)(U(x_2) - U(x_0)).$$

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

$$217 \quad 2w(p) + (w(d+p)-w(p)) = 2w(g+p), \text{ and}$$

$$218 \quad \frac{w(p) + w(d+p)}{2} = w(g+p)$$

219 follows, as required.  $\square$

220

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

222

223

224

225

226

227

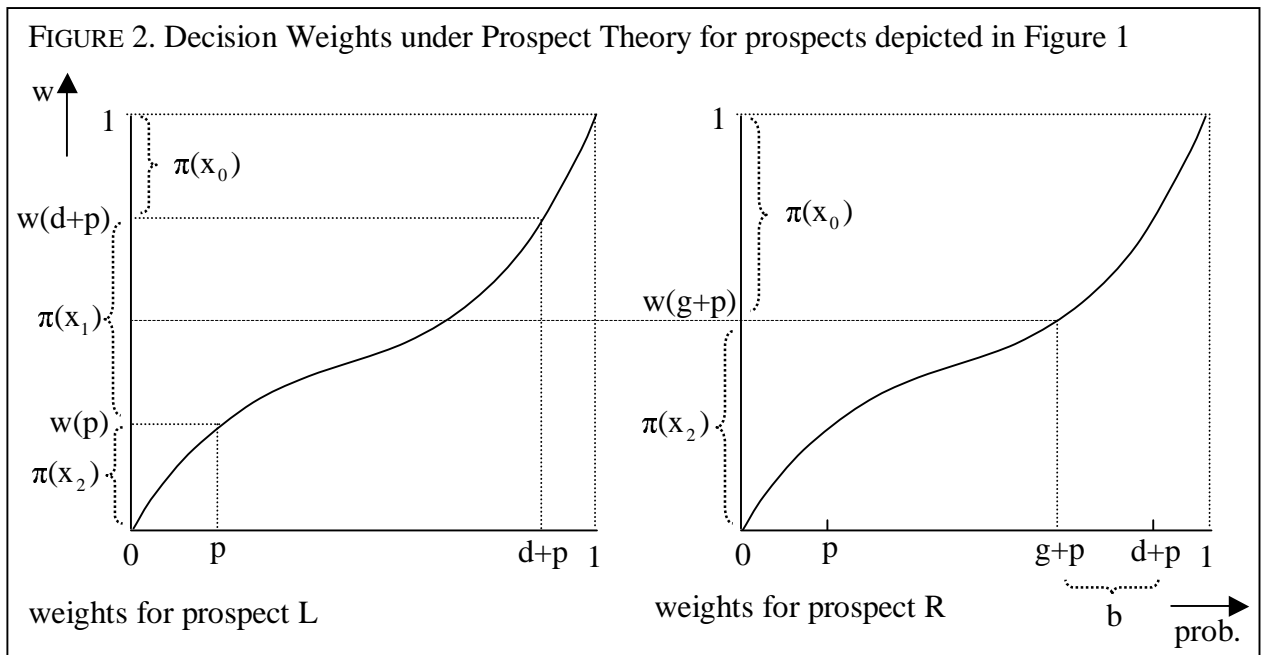
228

229

230

231

232



233

234 Our measurement technique is general in the sense that the weight-midpoint between

235 any two probabilities can be measured directly. The only richness of outcomes

236 needed is that there are three outcomes as above, i.e. equally spaced in utility units.

237

238

239

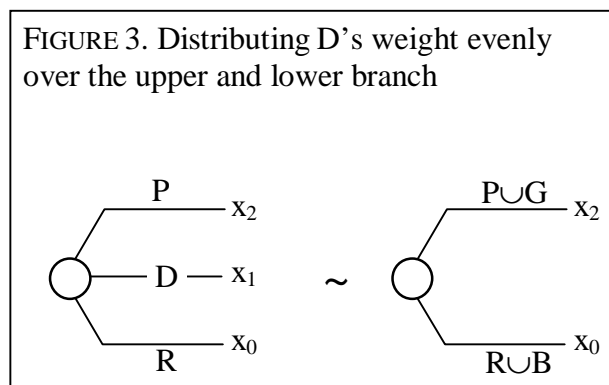
240

241

242

243

244



245 Our technique can readily be extended to Schmeidler's (1989) Choquet  
 246 expected utility, and prospect theory, for the case of events with unknown  
 247 probabilities. Figure 3 results from Figure 1 by replacing the probabilities  $p$ ,  $d$ , and  $r$   
 248 by exhaustive and mutually exclusive events  $P$ ,  $D$ , and  $R$ , replacing probability  $g$  by a  
 249 subevent  $G$  of  $R$ , and replacing probability  $b$  by event  $B = D - G$ . The indifference in  
 250 Figure 3 implies that the decision weight of event  $G$  captures half the decision weight  
 251 of event  $D$ , so that event  $G \cup P$  is the weight-midpoint between events  $P$  and  $D \cup P$ ,  
 252 similarly as above. Such a midpoint event  $G$  exists for all events  $P$  and  $D \cup P$  if the  
 253 event space is a continuum or at least is sufficiently rich, as for instance in Gilboa's  
 254 (1987) preference foundation.

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

## 262 4. THE EXPERIMENT: METHOD

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

265

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

270

271 *Stimuli; general.* Participants were asked two practice choice questions to familiarize  
 272 them with the experimental procedures. The choice questions involved simple  
 273 choices between two prospects named prospect L (left) and prospect R (right). Both  
 274 prospects yielded prizes depending on the outcome of a roll with two ten-sided dice,

275 generating probabilities  $j/100$ .<sup>1</sup> Prospects were framed as in Figure 4. Participants  
 276 were asked to indicate their choice by simply clicking on the button representing their  
 277 preferred prospect and were encouraged to answer the choice questions at their own  
 278 pace. In order to avoid potential confounding effects resulting from connotations with  
 279 words such as lottery or gamble, we used the more neutral term prospect in the  
 280 instructions. The position of each of the two prospects was counterbalanced between  
 281 participants in order to avoid a potential confounding effect that might result from  
 282 individual preference for a particular position of a prospect.

283  
 284 FIGURE 4. The framing of the prospect pairs

| PROSPECT L   |               |                | PROSPECT R   |               |            |
|--------------|---------------|----------------|--------------|---------------|------------|
| roll         | probability   | prize          | roll         | probability   | prize      |
| 1 till p     | p %           | $x_{i-1}$ euro | 1 till p     | p %           | $x_i$ euro |
| p+1 till 100 | $(100 - p)\%$ | Y euro         | p+1 till 100 | $(100 - p)\%$ | y euro     |

289  
 290  
 291 *Stimuli of the part measuring utility.* In the “outcome part” of the experiment, we set  
 292  $x_0 = 60$  and obtained values  $x_1$  and  $x_2$  to generate indifferences  $(0.25:x_1, 30) \sim$   
 293  $(0.25:60, 40)$  and  $(0.25:x_2, 30) \sim (0.25:x_1, 40)$ . Then  $x_1$  is the utility midpoint between  
 294  $x_0$  and  $x_2$  (Section 2). Because all further measurements in the experiment depended  
 295 on the values  $x_1$  and  $x_2$ , these values were elicited twice and the average of the two  
 296 values obtained was used as input in the rest of the experiment, so as to reduce noise.  
 297 To obtain indifferences we used a bisection choice method. This method, while time-  
 298 consuming, has been found to give more consistent results (Bostic, Herrnstein & Luce  
 299 1990) than direct matching.

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

<sup>1</sup> 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.

305 the  $j+1^{\text{th}}$  indifference interval from the  $j^{\text{th}}$  indifference interval  $[\ell^j, u^j]$ , we elicited  
 306 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  
 307 observed the choice between  $(0.25: \frac{\ell^j + u^j}{2}, 30)$  and  $(0.25: x_0, 40)$ . A left choice means  
 308 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  
 309 defined as the  $j+1^{\text{th}}$  indifference interval  $[\ell^{j+1}, u^{j+1}]$ . A right choice means that the  
 310 midpoint is smaller than  $x_1$ , so that  $x_1$  is contained in  $[\frac{\ell^j + u^j}{2}, u^j]$ , which was then  
 311 defined as the  $j+1^{\text{th}}$  indifference interval  $[\ell^{j+1}, u^{j+1}]$ . We did five iteration steps like  
 312 this, ending up with  $[\ell^6, u^6]$  (of length  $96 \times 2^{-5} = 3$ ), and took its midpoint as the  
 313 elicited indifference value  $x_1$ . We similarly elicited  $x_2$  (substitute  $x_2$  for  $x_1$  and  $x_1$  for  
 314  $x_0$  above).

315

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

324



354 To construct the  $j+1^{\text{th}}$  indifference interval  $[\ell^{j+1}, u^{j+1}]$  from the  $j^{\text{th}}$  indifference  
 355 interval  $[\ell^j, u^j]$ , we elicited whether the midpoint of  $[\ell^j, u^j]$  is larger or smaller than  
 356  $p+g$ . 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  
 357 right choice means that the midpoint is larger than  $g+p$ , so that  $g+p$  is contained in  
 358  $[\ell^j, \frac{\ell^j+u^j}{2}]$ , which was then defined as the  $j+1^{\text{th}}$  indifference interval  $[\ell^{j+1}, u^{j+1}]$ . A left  
 359 choice means that the midpoint is smaller than  $g+p$ , so that  $g+p$  is contained in  $[\frac{\ell^j+u^j}{2}$   
 360  $, u^j]$ , which was then defined as the  $j+1^{\text{th}}$  indifference interval  $[\ell^{j+1}, u^{j+1}]$ . We did five  
 361 iteration steps like this, ending up with  $[\ell^6, u^6]$ , and took its midpoint as the elicited  
 362 indifference probability  $p+g$ .

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

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

380

380

381

382

383

384

385

386

387

388

389

390

391

392

393

394

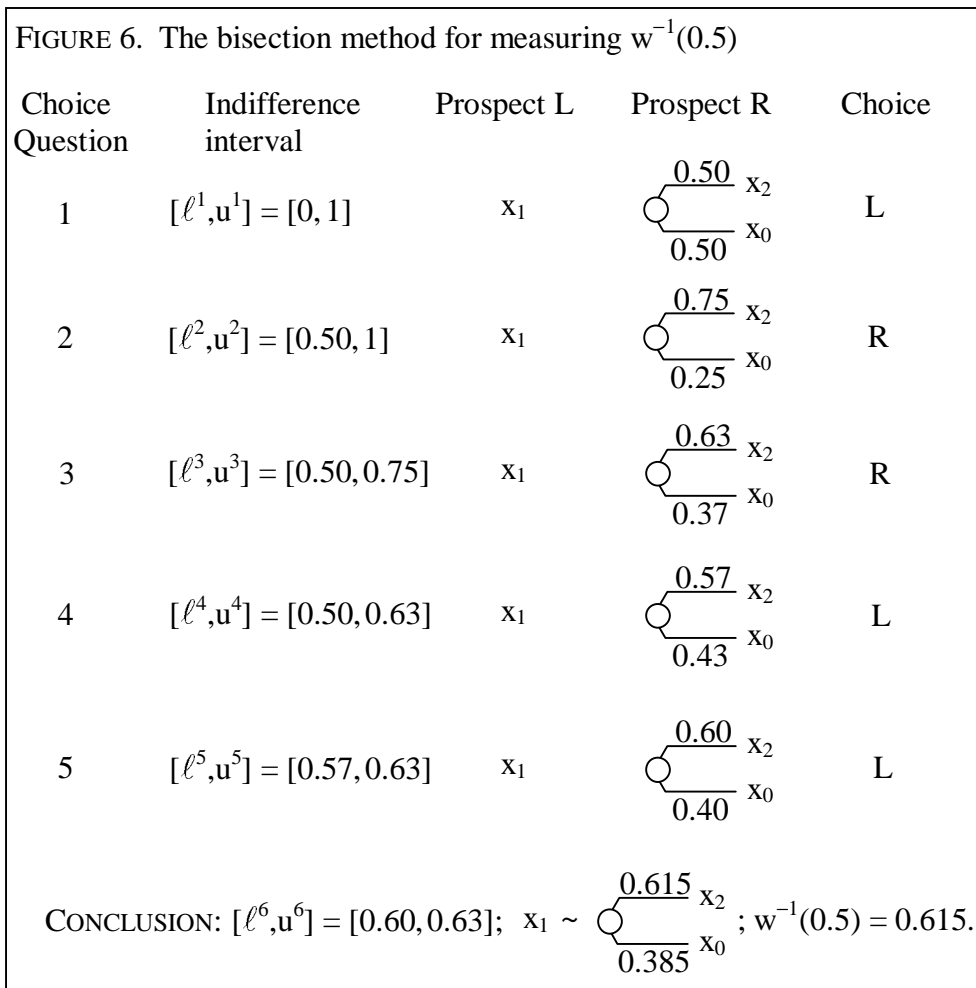
395

396

397

398

399



400

401

402

403

404

405

406

407

408

409

410

411

412

*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

413 the experiment. The average payment under real play was €77.57, so that the total  
 414 reward per participant was approximately  $10/11 \times 5 + 1/11 \times 77.57 = €11.60$ , while it  
 415 took participants about 20 minutes to complete the experiment.

416

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

## 424 5. THE EXPERIMENT: RESULTS

425 14 participants were excluded from the analysis because they gave erratic or  
 426 heuristic answers such as always choosing the left option or always choosing the right  
 427 option. The practice choices of this experiment also served to detect such erratic and  
 428 heuristic answers. These participants apparently did not understand the choices or did  
 429 not seriously think about them. Including these participants in the analysis would not  
 430 alter the results presented hereafter.

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

436

### 437 5.1. The Utility Function

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

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

447

## 448 **5.2. The Probability Weighting Function**

449

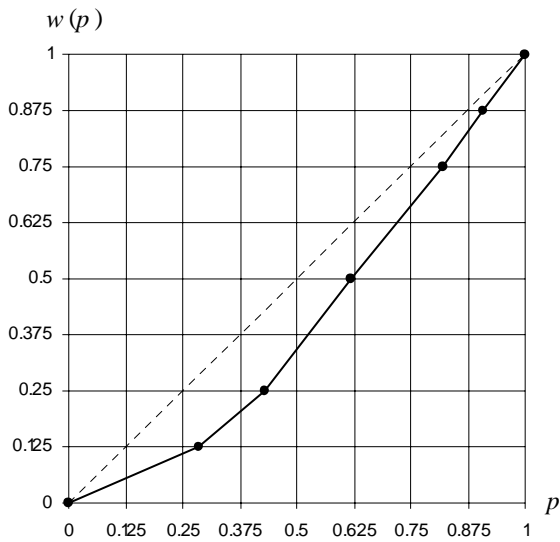
### 450 *5.2.1. Non-Parametric Analysis*

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

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

467

467 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              |

468

469

470

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    |

471

472

473

474

475

\*\*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%             |

476

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

479

### 480 5.2.2. Parametric Analysis

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

489

490 TABLE 3. Parameter Estimates

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

491 Standard errors are in parentheses.

492

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

$$495 \quad \sum_{i=1}^5 (w_i - \hat{w}_i)^2. \quad (5.1)$$

496 Here  $w_i$  is the  $i$ -th element of the sequence of the probability weights for which the  
 497 probabilities were elicited and  $\hat{w}_i$  is the  $i$ -th element of the estimated sequence of  
 498 probability weights under the various parametric specifications. For example, for  $w_1$   
 499  $= 0.125$ , we define  $p_1$  as the probability elicited in the experiment as  $w^{-1}(0.125)$  for  
 500 this subject, and then  $\hat{w}_1$  is the value that the probability weighting function of the  
 501 parametric family assigns to  $p_1$ . Its distance from 0.125 indicates how far the family  
 502 is from the data.

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

509

510

511

512

513

514

515

516

517

518

519

520

521

522

523

524

525

526

527

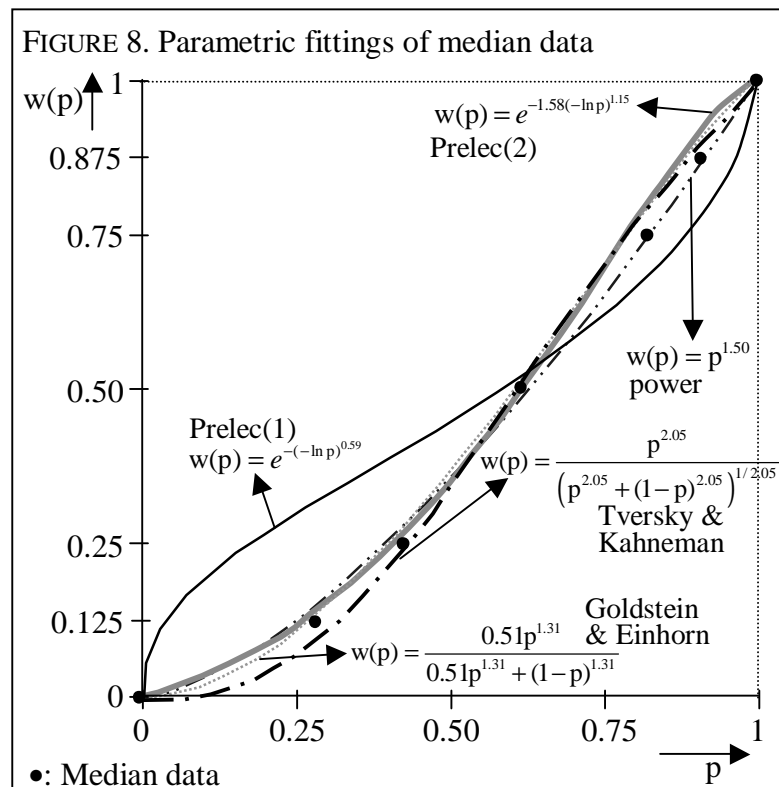


Figure 8 displays a remarkable variety of patterns, which should caution against the unqualified use of parametric fitting. The one-parameter family advocated



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

559 His method holds under prospect theory if we add the requirement that  $x_1 > x_2, x_1 >$   
560  $z_2, z_1 > z_2, \text{ and } z_1 > x_2.$

561 Ghirardato et al. (Definition 4) proposed another method to derive utility  
562 midpoints endogenously from preferences. They showed that  $\beta$  is the utility midpoint  
563 between  $\alpha$  and  $\gamma$  under prospect theory if the following indifferences hold:

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

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

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

583 To avoid lying to participants, we carefully formulated our instructions (end of  
584 Appendix A) as regards chaining. Approaches where participants are not told lies but  
585 do not receive complete information about the whole setup of the experiment, are not  
586 uncommon in experimental economics. For instance, in Holt & Laury (2002),  
587 participants were told that they would be paid for questions in a first part of the  
588 experiment (the “small-payment treatment”), which was true in a literal sense. They

589 were, however, not told that afterwards, in case they would continue to the second  
590 part of the experiment with a large-payment treatment, an event that was obvious to  
591 happen, their payments of the first part would be taken away. Thus, it was obvious  
592 beforehand that participants would end up not being paid for the first part of the  
593 experiment, but they were not informed about that. Empirical studies have found that  
594 weighting functions are mostly convex or inverse-S shaped, with the latter shape  
595 prevailing. In our study, the former shape was prevailing. This finding suggests once  
596 more that probability weighting is a volatile phenomenon, with results depending on  
597 framing and ways of measurement, and no phenomena holding in great generality.  
598

599

## 7. CONCLUSION

600 We have introduced a new way to measure probability weighting functions of  
601 rank-dependent utility (Gilboa 1987; Quiggin 1981; Schmeidler 1989) and prospect  
602 theory (Tversky & Kahneman 1992), both for risk and for uncertainty, and  
603 implemented it in an experiment for risk. Our results suggest that the weighting  
604 function may be more volatile and dependent on framing and measurement techniques  
605 than thought before. In particular, the result of parametric fitting can depend much on  
606 the parametric family chosen. The experiment demonstrated that our method is  
607 feasible and more efficient than methods used before.  
608

## 608 APPENDIX A. EXPERIMENTAL INSTRUCTIONS

609 [Instructions are translated from Dutch]

610

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

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

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

623

624

625

626

| 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  |

627

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

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

636 For your participation in this experiment, you receive 5 Euro anyhow.

637 Thereby, one participant of this experiment will be selected at random at the end of  
 638 this experiment. For this purpose, each participant randomly picks a sealed envelope



- 669 Chateauneuf, Alain (1991), "On the Use of Capacities in Modeling Uncertainty  
670 Aversion and Risk Aversion," *Journal of Mathematical Economics* 20, 343–369.
- 671 Chew, Soo Hong (1983), "A Generalization of the Quasilinear Mean with  
672 Applications to the Measurement of Income Inequality and Decision Theory  
673 Resolving the Allais Paradox," *Econometrica* 51, 1065–1092.
- 674 Chew, Soo Hong & Guofu Tan (2005), "The Market for Sweepstakes," *Review of*  
675 *Economic Studies* 72, 1009–1029.
- 676 Dekel, Eddie (1986), "An Axiomatic Characterization of Preferences under  
677 Uncertainty: Weakening the Independence Axiom," *Journal of Economic Theory*  
678 40, 304–318.
- 679 Donkers, A.C.D., Bertrand Melenberg, & Arthur H.O. van Soest (2001), "Estimating  
680 Risk Attitudes Using Lotteries; A Large Sample Approach," *Journal of Risk and*  
681 *Uncertainty* 22, 165–195.
- 682 Dybvig, Philip & Heracles M. Polemarchakis (1981), "Recovering Cardinal Utility,"  
683 *Review of Economic Studies* 48, 159–166.
- 684 Etchart, Nathalie (2004), "Is Probability Weighting Sensitive to the Magnitude of  
685 Consequences? An Experimental Investigation on Losses," *Journal of Risk and*  
686 *Uncertainty* 28, 217–235.
- 687 Ghirardato, Paolo, Fabio Maccheroni, Massimo Marinacci, & Marciano Siniscalchi  
688 (2003), "A Subjective Spin on Roulette Wheels," *Econometrica* 71, 1897–1908.
- 689 Gilboa, Itzhak (1987), "Expected Utility with Purely Subjective Non-Additive  
690 Probabilities," *Journal of Mathematical Economics* 16, 65–88.
- 691 Gilboa, Itzhak & David Schmeidler (1989), "Maxmin Expected Utility with a Non-  
692 Unique Prior," *Journal of Mathematical Economics* 18, 141–153.
- 693 Gold, Marthe R., Joanna E. Siegel, Louise B. Russell, & Milton C. Weinstein (1996),  
694 "Cost-Effectiveness in Health and Medicine." Oxford University Press, New  
695 York.
- 696 Goldstein, William M. & Hillel J. Einhorn (1987), "Expression Theory and the  
697 Preference Reversal Phenomena," *Psychological Review* 94, 236–254.
- 698 Gonzalez, Richard & George Wu (1999), "On the Shape of the Probability Weighting  
699 Function," *Cognitive Psychology* 38, 129–166.
- 700 Gul, Faruk (1991), "A Theory of Disappointment Aversion," *Econometrica* 59,  
701 667–686.

- 702 Harrison, Glenn W. (1986), "An Experimental Test for Risk Aversion," *Economic*  
703 *Letters* 21, 7–11.
- 704 Hey, John D. & Chris Orme (1994), "Investigating Generalizations of Expected  
705 Utility Theory Using Experimental Data," *Econometrica* 62, 1291–1326.
- 706 Holt, Charles A. & Susan K. Laury (2002), "Risk Aversion and Incentive Effects,"  
707 *American Economic Review* 92, 1644–1655.
- 708 Jullien, Bruno & Bernard Salanié (2000), "Estimating Preferences under Risk: The  
709 Case of Racetrack Bettors," *Journal of Political Economy* 108, 503–530.
- 710 Keeney, Ralph L. & Howard Raiffa (1976), "*Decisions with Multiple Objectives.*"  
711 Wiley, New York (Second edition 1993, Cambridge University Press,  
712 Cambridge, UK).
- 713 Köbberling, Veronika & Peter P. Wakker (2003), "Preference Foundations for  
714 Nonexpected Utility: A Generalized and Simplified Technique," *Mathematics of*  
715 *Operations Research* 28, 395–423.
- 716 Loomes, Graham & Robert Sugden (1982), "Regret Theory: An Alternative Theory of  
717 Rational Choice under Uncertainty," *Economic Journal* 92, 805–824.
- 718 Mukerji, Sujoy & Jean-Marc Tallon (2001), "Ambiguity Aversion and  
719 Incompleteness of Financial Markets," *Review of Economic Studies* 68, 883–904.
- 720 Prelec, Drazen (1998), "The Probability Weighting Function," *Econometrica* 66,  
721 497–527.
- 722 Quiggin, John (1981), "Risk Perception and Risk Aversion among Australian  
723 Farmers," *Australian Journal of Agricultural Economics* 25, 160–169.
- 724 Rabin, Matthew (2000), "Risk Aversion and Expected-utility Theory: A Calibration  
725 Theorem," *Econometrica* 68, 1281–1292.
- 726 Schmeidler, David (1989), "Subjective Probability and Expected Utility without  
727 Additivity," *Econometrica* 57, 571–587.
- 728 Starmer, Chris (1992), "Testing New Theories of Choice under Uncertainty Using the  
729 Common Consequence Effect," *Review of Economic Studies* 59, 813–830.
- 730 Tversky, Amos & Craig R. Fox (1995), "Weighing Risk and Uncertainty,"  
731 *Psychological Review* 102, 269–283.
- 732 Tversky, Amos & Daniel Kahneman (1981), "The Framing of Decisions and the  
733 Psychology of Choice," *Science* 211, 453–458.

- 734 Tversky, Amos & Daniel Kahneman (1992), “Advances in Prospect Theory:  
735 Cumulative Representation of Uncertainty,” *Journal of Risk and Uncertainty* 5,  
736 297–323.
- 737 Vind, Karl (1991), “Independent Preferences,” *Journal of Mathematical Economics*  
738 20, 119–135.
- 739 Vind, Karl (2003), “*Independence, Additivity, Uncertainty.*” With contributions by B.  
740 Grodal. Springer, Berlin. First version 1969.
- 741 Wakker, Peter P. & Daniel Deneffe (1996), “Eliciting von Neumann-Morgenstern  
742 Utilities when Probabilities Are Distorted or Unknown,” *Management Science*  
743 42, 1131–1150.
- 744 Wald, Abraham (1950), “*Statistical Decision Functions.*” Wiley, New York.
- 745 Young, H. Peyton (1990), “Progressive Taxation and Equal Sacrifice,” *American*  
746 *Economic Review* 80, 253–266.
- 747