

---

Parameter-Free Elicitation of Utility and Probability Weighting Functions

Author(s): Mohammed Abdellaoui

Source: *Management Science*, Vol. 46, No. 11 (Nov., 2000), pp. 1497-1512

Published by: [INFORMS](#)

Stable URL: <http://www.jstor.org/stable/2661664>

Accessed: 25/03/2013 01:45

---

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at  
<http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



INFORMS is collaborating with JSTOR to digitize, preserve and extend access to *Management Science*.

<http://www.jstor.org>

# Parameter-Free Elicitation of Utility and Probability Weighting Functions

Mohammed Abdellaoui

CNRS, GRID, ENS de Cachan, Department of Economics and Management, 61 avenue du Président Wilson,  
94235 Cachan, France  
abdellaoui@grid.ens-cachan.fr

This paper proposes a two-step method to successively elicit utility functions and decision weights under rank-dependent expected utility theory and its "more descriptive" version: cumulative prospect theory. The novelty of the method is that it is parameter-free, and thus elicits the whole individual preference functional without imposing any prior restriction. This method is used in an experimental study to elicit individual utility and probability weighting functions for monetary outcomes in the gain and loss domains. Concave utility functions are obtained for gains and convex utility functions for losses. The elicited weighting functions satisfy upper and lower subadditivity and are consistent with previous parametric estimations. The data also show that the probability weighting function for losses is more "elevated" than for gains.

*(Decision Making; Expected Utility; Rank-Dependent Expected Utility; Cumulative Prospect Theory; Probability Weighting Function)*

## 1. Introduction

In a seminal paper, Kahneman and Tversky (1979) present experimental evidence that preferences between risky prospects are not linear in probabilities. They propose, as well, a theory of choice under risk, Prospect Theory (PT), suggesting that a probability weighting function (that maps the unit interval into itself with discontinuities at 0 and 1) exhibiting overweighting of small probabilities and underweighting of moderate and high probabilities may explain the observed nonlinearities. Subsequently, modern generalizations of PT were proposed through Rank-Dependent Expected Utility (RDEU) theory (Quiggin 1982, Wakker 1994), and Cumulative Prospect Theory (CPT) (Tversky and Kahneman 1992). Their main characteristic, in decision under risk, consists of allowing not only the transformation of outcomes into utilities, but also the transformation of decumulative probabilities to obtain decision weights through a probability weighting function. This innovation, however, has been perceived as a factor complicating utility mea-

surement, and therefore, the elicitation of RDEU and CPT models.

A variety of methods have been used to determine the shapes of the utility function and the probability weighting function under RDEU and CPT. The predominant approach prespecifies parametric forms for these functions and then estimates them through standard techniques (e.g., Tversky and Kahneman 1992, Camerer and Ho 1994, Hey and Orme 1994, Tversky and Fox 1995). However, assuming specific functional forms for the utility function and the probability weighting function makes inference about the shapes of these functions dependent on the choice of functional forms.

Two research strategies can avoid the potential problems of parametric estimation. The first strategy consists of testing simple preference conditions to obtain information about the shape of either the utility function or the probability weighting function. Wu and Gonzalez (1996, 1998) follow this strategy. They find that aggregate behavior is consistent with a

concave-convex shaped weighting function independent of any assumption about the value function. The second strategy consists of eliciting the utility and probability weighting functions at the level of individuals, without any parametric assumption. This approach is more demanding, but, in return, provides direct measurements of both functions. This paper implements the latter strategy, proposing a two-step method to successively elicit the decision-maker's utility and weighting functions under CPT.

The first step consists of constructing a sequence of outcomes equally spaced in utility by means of the trade-off method initially proposed by Wakker and Deneffe (1996). Contrary to the other existing methods of utility elicitation, the trade-off method is "robust" against probability distortion. The second step uses the sequence of outcomes to obtain a sequence of probabilities equally spaced in terms of probability weighting. These two steps were computerized to allow a comfortable questioning of subjects and to avoid some expected framing effects by means of a special sequencing of choice questions. This proposed parameter-free elicitation method is used in an experimental investigation to study the shapes of the utility function and the probability weighting function for gains and losses.

Although Tversky and Kahneman (1992, p.311), consider the estimation of such a complex model as CPT to be problematic; this paper shows that it can be done without much trouble. Many experimental findings made in Tversky and Kahneman (1992) and Wu and Gonzalez (1996) have been confirmed in this paper by means of parameter-free elicitation of the utility and probability weighting functions. To my knowledge, this is the first paper that entirely elicits the CPT model (for both gains and losses) at the level of individual subjects without any parametric assumption.<sup>1</sup>

<sup>1</sup> A few months after a first version of this paper was completed, Bleichrodt and Pinto (1998) and Gonzalez and Wu (1999) finished two papers proposing two methods to elicit probability weighting functions. The first paper reports experimental results regarding probability weighting in medical decision making (using the tradeoff method). The second investigates individual differences in probability weighting for monetary gains through an alternating least square estimation method.

Given that the elicitation of the probability weighting function in this paper needs the construction of the utility function to be carried out first, the first empirical question addressed concerns the shape of the utility function. Its qualitative properties issuing from the psychological principle of diminishing sensitivity are confirmed here, in agreement with other recent findings (Wakker and Deneffe 1996, Fennema and van Assen 1998, Fox and Tversky 1998). Then the question of the shape of the probability weighting function is addressed. The data confirm that individuals transform probabilities consistently with the psychological principle of diminishing sensitivity, with the two end points of the probability interval serving as reference points. Overall, these results are consistent (for gains) with those obtained recently by Tversky and Fox (1995).

This paper also elicits probability weighting functions for losses. This allows a straightforward comparison of the treatment of probabilities for gains and losses at the level of individual subjects. Indeed, the data confirm the existence of a significant difference between the probability weighting function for gains and the probability weighting function for losses. Moreover, the data suggest a descriptive superiority of CPT over RDEU.

Finally, the hypothesis of linearity of the probability weighting function for moderate probabilities is investigated. This question has received rather contradictory answers in the experimental literature. Camerer (1992), Harless and Camerer (1994), and Abdellaoui and Munier (1998), for instance, obtained results through nonparametric techniques suggesting a linear weighting function for intermediate probabilities (see also Cohen and Jaffray 1988). On the contrary, Wu and Gonzalez (1996), among others, found support for nonlinearity. This paper confirms the latter (i.e., nonlinearity).

The paper proceeds as follows. Section 2 reviews RDEU theory and CPT. Then, some previous experimental studies of parametric estimation of RDEU or CPT preference functionals are presented. Section 3 describes the two-step elicitation method. Section 4 describes an experimental elicitation of utility functions and weighting functions for gains and losses. The results of this experiment are given in §5. Finally, §6

summarizes and discusses the experimental findings presented in §5.

## 2. Review of RDEU and CPT

Let  $X$  be a set of *consequences*, e.g., monetary outcomes. We assume that  $X$  includes a *neutral outcome* denoted 0. Positive and negative numbers are interpreted as *gains* and *losses* respectively. In decision under risk a *prospect*, or *lottery*, is described by a finite probability distribution over  $X$ . Thus  $(x_1, p_1; \dots; x_n, p_n)$  is the prospect yielding outcome  $x_i$  with probability  $p_i$ ,  $i=1, \dots, n$ .

According to RDEU theory, the "utility" of a lottery  $P = (x_1, p_1; \dots; x_n, p_n)$  in which  $x_1 \leq \dots \leq x_n$ , denoted  $V_{RDEU}(P)$ , depends upon a utility function  $u$  and a probability weighting function  $w$ . The function  $u$  is a strictly increasing function over  $X$  and the function  $w$  is a strictly increasing function from  $[0, 1]$  to  $[0, 1]$  with  $w(0)=0$  and  $w(1)=1$ .  $V_{RDEU}(P)$  is given by:

$$V_{RDEU}(P) = \sum_{i=1}^n \pi_i u(x_i) \quad (1)$$

where  $\pi_i = w(\sum_{k=i}^n p_k) - w(\sum_{k=i+1}^n p_k)$  for  $i \leq n-1$  and  $\pi_n = w(p_n)$ . Note that the decision weights  $\pi_i$  sum to one and depend on the ranking of outcomes  $x_i$ ,  $i=1, \dots, n$ . When there is no probability weighting, i.e.,  $w(p)=p$ , for all  $p \in (0, 1)$ , RDEU reduces to Expected Utility (EU).

CPT is a "more descriptive" and more general representation of individual preferences under risk than RDEU. It postulates that the carriers of utility are gains and losses, not the final asset position as typically assumed in EU and RDEU. Therefore, it assumes a continuous strictly increasing utility function, denoted  $u$  and satisfying  $u(0)=0$ . Furthermore, this theory invokes two weighting functions, denoted by  $w^+$  and  $w^-$ , for gains and losses respectively, leading to "sign dependence."

Following CPT, the "utility" of a prospect  $P = (x_1, p_1; \dots; x_n, p_n)$  in which  $x_1 \leq \dots \leq x_r \leq 0 \leq x_{r+1} \leq \dots \leq x_n$  is:

$$V_{CPT}(P) = \sum_{i=1}^r \pi_i^- u(x_i) + \sum_{j=r+1}^n \pi_j^+ u(x_j) \quad (2)$$

where the decision weights for losses  $\pi_i^-$  and the decision weights for gains  $\pi_j^+$  are defined by:

Table 1 Some "Typical" Parametric Specifications of  $u$  and  $w$

Experiments	Models	$u(x)$	$w(p)$
Currim and Sarin (1989)	PT	$\frac{1-e^{-\alpha x}}{1-e^{-\alpha}}$	$\frac{1-e^{-\beta p}}{1-e^{-\beta}}$ $a + bp + cp^2$ $\log(a + be^{c \log p})$
Tversky and Kahneman (1992)	CPT	$\begin{cases} x^\alpha & \text{if } x \geq 0 \\ -\lambda(-x)^\beta & \text{if } x < 0 \end{cases}$	$\frac{p^\gamma}{[p^\gamma + (1-p)^\gamma]^\frac{1}{\gamma}}$
Tversky and Fox (1995)	CPT	$x^\alpha$	$\frac{\delta p^\gamma}{[\delta p^\gamma + (1-p)^\gamma]^\frac{1}{\gamma}}$ <sup>(1)</sup>
Wu and Gonzalez (1996)	PT CPT	$x^\alpha$	$\frac{\delta p^\gamma}{[\delta p^\gamma + (1-p)^\gamma]^\frac{1}{\gamma}}$ $\frac{p^\gamma}{[p^\gamma + (1-p)^\gamma]^\frac{1}{\gamma}}$ $a + bp$ $\exp(-(-\ln p)^\omega)$ <sup>(2)</sup>

<sup>(1)</sup> Used by Goldstein and Einhorn (1987) and Lattimore et al. (1992).

<sup>(2)</sup> Initially proposed by Prelec (1998).

$\pi_i^- = w^-(\sum_{k=i}^r p_k) - w^-(\sum_{k=i+1}^r p_k)$  for  $i \geq 2$  and  $\pi_1^- = w^-(p_1)$ ;  $\pi_j^+ = w^+(\sum_{k=j}^n p_k) - w^+(\sum_{k=j+1}^n p_k)$  for  $j \leq n-1$  and  $\pi_n^+ = w^+(p_n)$ . Note that these weights do not necessarily sum to one, and that RDEU corresponds to the special case where the weighting function for losses is the *dual* of the weighting function for gains, i.e.,  $w^-(p) = 1 - w^+(1-p)$  for all  $p \in (0, 1)$ .

Up to now, most experimental studies used parametric specifications to infer the shapes of utility functions and weighting functions from individual choices. Tversky and Kahneman (1992) and Tversky and Fox (1995) gave parametric estimates for the median subject. In Camerer and Ho (1994) all subjects were assumed to have the same underlying preferences. Other studies estimate parametric forms at the level of individual subjects, e.g., Lattimore et al. (1992), Hey and Orme (1994), Abdellaoui (1995a). Table 1 gives some typical parametric specifications of  $u$  and  $w$ . The results of these experimental studies are mostly consistent with an inverse S-shaped weighting function, concave for small probabilities and convex for moderate and high probabilities. They are also consistent with a concave utility function for gains.

More recently, Wu and Gonzalez (1996) used a qualitative approach to analyze the shape of the probability weighting function. They proposed two common-consequence conditions that are necessary and sufficient for concavity and convexity of the weighting function. These conditions were used to test the empirical properties of the weighting function at an aggregate level. Their tests validate the basic inverse S-shape and suggest that the nonlinearity of the probability weighting function (for gains) is not due merely to boundary effects. The authors also used their aggregated data to give parametric estimates of the probability weighting function (see Table 1).

While recognizing the merits of the parametric studies, one must also agree that the findings may have been confounded by the particular parametric families chosen. In light of this issue, nonparametric elicitation of utility and weighting functions can allow reliable tests of the key features regarding the shapes of these functions. More specifically, if done properly, nonparametric elicitation provides a way to assess the deviation of the data from parametric specifications.

### 3. Elicitation of Utility and Weighting Functions

This section proposes a two-step method to elicit utility and weighting functions. In the first step, a *standard sequence* of outcomes, i.e., equally spaced outcomes in terms of utility, is constructed. Then, in the second step, this standard sequence is used in simple risky choices to obtain a standard sequence of probabilities, i.e., equally spaced probabilities in terms of the weighting function.

#### 3.1. Standard Sequences of Outcomes

As pointed out by Wakker and Deneffe (1996), the two common methods of utility elicitation, the certainty equivalent (CE) method and the probability equivalent (PE) method, suffer from an important cause of violation of expected utility theory: the "psychological transformation of probabilities". The trade-off method proposed by these two authors has the advantage of minimizing the role of probabilities in the assessment

process. Moreover, it allows constructing standard sequences of outcomes even when probabilities are distorted or unknown.<sup>2</sup>

A standard sequence of positive outcomes (i.e., gains), is constructed as follows. An outcome  $x_1$  is determined to make the subject indifferent between the prospects  $(x_0, p; R, 1-p)$  and  $(x_1, p; r, 1-p)$ , denoted  $(x_0, p; R)$  and  $(x_1, p; r)$  respectively, where  $0 \leq r < R < x_0 < x_1$ , and  $p \in (0,1)$ ;  $r$ ,  $R$  and  $x_0$  are held fixed. This means that the trade-off of  $R$  for  $r$ , over the "(1-p)-axis," outweighs the trade-off of  $x_0$  for  $x_1$ , over the '(p)-axis' (see Wakker 1989, p. 35). Then, an outcome  $x_2$  is determined to make the subject indifferent between the prospects  $(x_1, p; R)$  and  $(x_2, p; r)$ . Thus, under CPT, the two constructed indifferences give the following two equations:

$$\begin{aligned} w^+(p)u(x_0) + (1 - w^+(p))u(R) \\ = w^+(p)u(x_1) + (1 - w^+(p))u(r), \end{aligned} \quad (3)$$

$$\begin{aligned} w^+(p)u(x_1) + (1 - w^+(p))u(R) \\ = w^+(p)u(x_2) + (1 - w^+(p))u(r). \end{aligned} \quad (4)$$

Together, these equations imply:

$$u(x_1) - u(x_0) = u(x_2) - u(x_1). \quad (5)$$

Equation (5) implies that  $x_1$  is a midpoint outcome in terms of utility in the sequence  $x_0, x_1, x_2$ , i.e., the trade-off of  $x_0$  for  $x_1$  is considered as equivalent to the trade-off of  $x_1$  for  $x_2$ , the trade-off of  $R$  for  $r$  being used as a "measuring rod." This conclusion holds also under EU, i.e.,  $w^+(p) = p$ . A standard sequence  $x_0, \dots, x_n$  needs the construction of  $n$  indifferences  $(x_{i-1}, p; R) \sim (x_i, p; r)$ ,  $i = 1, \dots, n$ . A similar process applied to the corresponding negative prospects allows us to obtain a standard sequence of losses.<sup>3</sup> Such a standard sequence is increasing for gains (i.e.,  $0 < x_0 < \dots < x_n$ ) and decreasing for losses (i.e.,  $x_n < \dots < x_0 < 0$ ).

<sup>2</sup>To our knowledge, Miyamoto (1988) was the first to suggest the idea of using preference conditions in non-EU contexts by restricting oneself to particular subsets of lotteries.

<sup>3</sup>In this case,  $0 \geq r > R > x_0 > x_1$ .



### 3.2. Standard Sequences of Probabilities

Suppose that  $x_0, \dots, x_n$  is a standard sequence of gains. Consider the probabilities  $p_i, i=1, \dots, n-1$ , satisfying:

$$(x_n, p_i; x_0) \sim (x_i, 1; x_0) \equiv x_i, \quad (6)$$

where  $0 < x_0 < x_n$ . Under CPT, this indifference implies

$$w^+(p_i) = \frac{u(x_i) - u(x_0)}{u(x_n) - u(x_0)}, \quad i=1, \dots, n-1. \quad (7)$$

Because  $u(x_i) - u(x_{i-1})$  is constant for  $i=1, \dots, n$ , the above equations become

$$w^+(p_i) = \frac{i}{n}, \quad i=1, \dots, n-1. \quad (8)$$

This means that the assessment of a standard sequence of outcomes and the construction of the indifferences (6) elicit the probability weighting function  $w^+$ .

When  $x_0, \dots, x_n$  is a (decreasing) standard sequence of losses, we can find a sequence of probabilities  $q_1, \dots, q_{n-1}$  satisfying  $(x_n, q_j; x_0) \sim x_j$ , where  $x_n < x_0 < 0$  and  $j=1, \dots, n-1$ . Similarly, under CPT, we have  $u(x_j) = w^-(q_j)u(x_n) + (1 - w^-(q_j))u(x_0)$ ,  $j=1, \dots, n-1$ , and therefore:

$$w^-(q_j) = \frac{j}{n}, \quad j=1, \dots, n-1 \quad (9)$$

for losses.

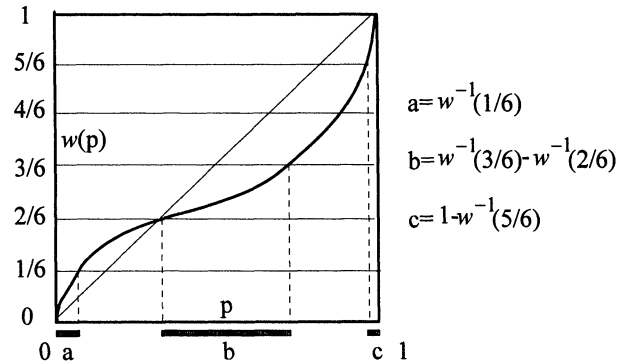
Indifferences (6) show that error in utility measurement can propagate towards  $p_i$  (and, therefore, toward the probability weighting function) through  $x_i$  and  $x_n$ . Thus the reliability of the measurement of probability weighting depends also (and mainly) on the reliability of the assessment of the standard sequence of outcomes (see §5.1).

To summarize, we may experimentally elicit the weighting function in two successive steps:

- (i) construction of a standard sequence of outcomes (trade-off experiments: TO-experiments);
- (ii) determination of a corresponding standard sequence of probabilities (probability weighting experiments: PW-experiments).

Preference conditions concerning the shape of the weighting function were formulated by Segal (1987), Tversky and Wakker (1995), Wu and Gonzalez (1996, 1998) and Prelec (1998). Tversky and Wakker (1995) define two properties of the weighting function that are rarely violated, at least in parametric estimation. These properties are called *lower subadditivity* (LS) and *upper*

Figure 1 Illustration of Subadditivity



*subadditivity* (US). LS implies that a 'lower interval'  $[0, q]$  has more impact than a middle interval  $[p, p+q]$ , and US means that an 'upper' interval  $[1-q, 1]$  has more impact than a middle interval  $[p, p+q]$ , provided the middle interval is bounded away from the lower end point 0 and upper end point 1. LS and US are generally referred to as the *possibility effect* and the *certainty effect* respectively. Formally, for constant  $\varepsilon \geq 0$ , a weighting function  $w$  satisfies LS (or the possibility effect) if:

$$w(q) - w(0) \geq w(p+q) - w(p) \quad \text{whenever } p+q \leq 1 - \varepsilon; \quad (10)$$

and for constant  $\varepsilon' \geq 0$ ,  $w$  satisfies US (or the certainty effect) if:

$$w(1) - w(1-q) \geq w(p+q) - w(p) \quad \text{whenever } p \geq \varepsilon'. \quad (11)$$

In the presence of a standard sequence of probabilities  $p_i = w^{-1}(\frac{i}{n})$ ,  $i=1, \dots, n-1$ , where  $w^{-1}$  is the *inverse* of the weighting function  $w$ , these inequalities can be tested rather straightforwardly. Figure 1 gives an illustration of subadditivity for  $n=6$ . Thus, Inequalities (10) and (11) correspond to Inequalities  $a \leq b$  and  $c \leq b$  respectively.

## 4. Experiments

### 4.1. Subjects and Procedure

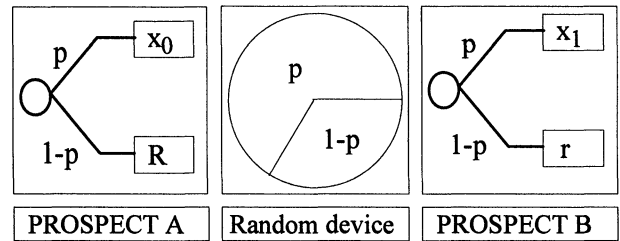
This section describes two experiments in which utility and weighting functions for gains and losses are elicited under CPT. Special efforts were made to obtain high-quality data. Sixty-four subjects were recruited to participate in these experiments. A group of 18 subjects participated in a pilot study which allowed the adjustment of the software and the experimental protocol. Forty-six subjects participated in the final study.<sup>4</sup>

All subjects were either undergraduate or Ph.D. students in economics. They were all acquainted with probabilities and expectations, and most of them had heard about expected utility at some point, but not shortly before the experiment. Each subject was paid FF 150 (approximately U.S. \$25) for participation.

As pointed out by Wakker and Deneffe (1996), to make the curvature of the utility function sufficiently pronounced, it is necessary to investigate a sufficiently wide interval of outcomes. For all the prospects used in the present experimental study, the outcomes (gains and losses) were between U.S. \$200 and U.S. \$4,000. Subjects were informed before the experiment that, for gains and only for gains, one subject would be randomly selected to play for actual money on one of her answers in TO-experiments and one of her answers in PW-experiments.<sup>5</sup> For losses, a similar lottery incentive mechanism is ethically objectionable. Hence, only hypothetical payoffs were used (e.g., Camerer 1989, Thaler and Johnson 1990). It will be seen later that the reliability of subjects in TO-experiments was higher for losses than for gains.

The experiments were conducted in small groups of six subjects in the experimental decision laboratory at the Ecole Normale Supérieure de Cachan, Department of Economics and Management. Subjects were seated in front of personal computers and were told to take their time, and encouraged to go at their own pace. Be-

Figure 2 Generic Display for TO-Experiments



fore the experiment started, approximately 10 minutes were devoted to explanations regarding the displayed problems of choice through "choice trials."

All the assessed quantities, outcomes, and probabilities were obtained through a series of choice questions. Each question consisted of an outright choice between two prospects. The construction of the standard sequence of outcomes was always carried out first because the trade-off answers were used as input in the elicitation of weighting functions. Thus, subjects participated in two successive and separate 40-minute sessions: one for gains and the other for losses. Each session consisted successively of TO-experiments and PW-experiments.

### 4.2. Methods and Stimuli for Utility Functions

The subjects were instructed to designate in each displayed pair of prospects which one they would prefer to be faced with. Figure 2 illustrates the generic display for the trade-off questions.

In the experiments involving gains, as in those involving losses, standard sequences of six outcomes:  $x_1, \dots, x_6$ , were constructed. For gains and losses, the outcomes  $|x_0|$ ,  $|r|$ , and  $|R|$  were fixed at the following levels: FF 1,000 (U.S. \$200), 0 and FF 500 (U.S. \$100) respectively. Subjects were not asked to find switching outcomes, but they were asked to express outright choices between prospects.<sup>6</sup> Bostic et al. (1990) report experimental evidence that choice is "more consistent" than matching. Indeed, they find that the bisection method for elicitation greatly reduces choice versus pricing reversals, and that judged CEs and

<sup>4</sup> Six subjects were discarded; due to a network problem, their data were lost.

<sup>5</sup> Subjects' answers were not collected as switching outcomes or probabilities (i.e., through matching), but as binary choices. Then, playing an answer for actual money consisted of playing the preferred prospect.

<sup>6</sup> Fennema and van Assen (1998) allowed their subjects to match switching outcomes.

**Table 2** Assessing  $x_1$  and  $p_1$  Through Bisection (For Gains)

Question	Alternatives <sup>†</sup>	Outcomes (F.F.) $x_1 \in$	Choice	Alternatives	Probabilities ( $\times 1\%$ ) $p_1 \in$	Choice
1	A = ( $x_0$ , $p$ ; R) B = ( <u>3,500</u> , $p$ ; r)	[1,000; 6,000]	A	A = ( $x_1$ , 1) B = ( $x_6$ , <u>50</u> ; $x_0$ )	[0, 100]	A
2	A = ( $x_0$ , $p$ ; R) B = ( <u>4,700</u> , $p$ ; r)	[3,500; 6,000]	A	A = ( $x_1$ , 1) B = ( $x_6$ , <u>75</u> ; $x_0$ )	[50, 100]	A
3	A = ( $x_0$ , $p$ ; R) B = ( <u>5,300</u> , $p$ ; r)	[4,700; 6,000]	B	A = ( $x_1$ , 1) B = ( $x_6$ , <u>87</u> ; $x_0$ )	[75, 100]	B
4	A = ( $x_0$ , $p$ ; R) B = ( <u>5,000</u> , $p$ ; r)	[4,700; 5,300]	A	A = ( $x_1$ , 1) B = ( $x_6$ , <u>81</u> ; $x_0$ )	[75, 87]	A
5	A = ( $x_0$ , $p$ ; R) B = ( <u>5,100</u> , $p$ ; r)	[5,000; 5,300]	A	A = ( $x_1$ , 1) B = ( $x_6$ , <u>84</u> ; $x_0$ )	[81, 87]	A
6				A = ( $x_1$ , 1) B = ( $x_6$ , <u>85</u> ; $x_0$ )	[84, 87]	A
End		[5,100; 5,300]			[85, 87]	

<sup>†</sup>:  $x_0$  = FF 1,000,  $p$  = 2/3,  $r$  = 0,  $R$  = FF 500.

choice-based CEs can differ substantially for some lotteries.

Five iterations (questions) were necessary to obtain each switching outcome  $x_i$ ,  $i=1, \dots, 6$ , through bisection. Suppose that  $x_{i-1}$  is a known gain. To determine gain  $x_i$ , the subject was asked in the  $j$ -th choice question to choose between prospects  $A = (x_{i-1}, p; 500)$  and  $B = (x_i^j, p; 0)$ , where  $x_i^j$  is taken as the midpoint of the interval of "feasible outcomes" corresponding to the  $j$ -th iteration (question). The interval corresponding to the first iteration was  $[x_{i-1}, x_{i-1} + \Delta]$ . Of course  $\Delta$  needs to be fixed at a level that guarantees a strong preference of prospect  $B = (x_{i-1} + \Delta, p; 0)$  over prospect  $A$  for all  $i=1, \dots, 6$ .  $|\Delta|$  was fixed at a level of FF 5,000 (U.S. \$900) for gains and losses.<sup>7</sup>

Table 2 gives the procedure followed by the computer program to determine  $x_1$  for a subject who chooses prospect  $A$  in iterations 1, 2, 4, and 5. Note that the process consists of narrowing the interval containing  $x_1$ . In the fifth iteration, this outcome ends up in the interval [5,100; 5,300]. Then, the computer program

chooses the midpoint value of this interval (FF 5,200) as  $x_1$ . Only multiples of 100 were admitted as "feasible" switching values. In fact, when a midpoint value of an interval is not a multiple of 100, the computer program screens/saves the closest multiple of 100 on the "left" of this value. Thus, in the third iteration, 5,300 was used as a midpoint value instead of 5,350.<sup>8</sup>

Contrary to Wakker and Deneffe (1996), who elicited utility functions without specifying the probability  $p$  (i.e., in a context of ambiguity),  $p$  was given the value of 2/3 in the present experiment. In Fennema and van Assen (1998) and Bleichrodt and Pinto (1998),  $p$  was given the values 1/3 and 1/2 respectively. All the recent experimental studies using the trade-off method, including the present experiment, produced very similar results.

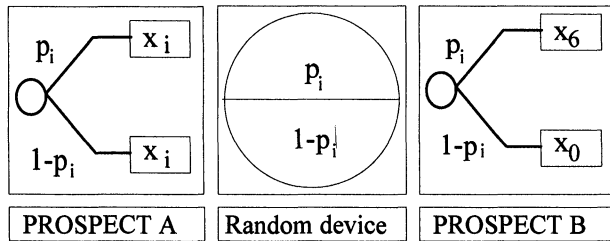
Once the standard sequence  $x_1, \dots, x_6$  is obtained, the computer program checks the subject's reliability by asking her to choose again between two prospects

<sup>8</sup> The procedure followed by the computer program to determine  $x_1$  for losses is similar to that described in Table 2, except that the choice pattern AABAA must be replaced by the choice pattern BBABB.

<sup>7</sup> This was inferred from the results of the pilot study.



Figure 3 Generic Display for PW-Experiments



corresponding to the fourth iteration for each  $x_i$ ,  $i=1, \dots, 6$  (see Table 2). Then, the next step consists of using these outcomes to elicit the probability weighting function.

### 4.3. Methods and Stimuli for Probability Weighting Functions

In the PW-experiments, each subject was asked a new series of choice questions aiming to determine the probabilities  $p_1, \dots, p_5$  that make her indifferent between the outcome  $x_i$  and the prospect  $(x_6, p_i; x_0)$ ,  $i=1, \dots, 5$  (see Figure 3). A sixth probability  $p'_2$  such that  $(x_3, 1) \sim (x_4, p'_2; x_2)$  was assessed to check the reliability of the elicitation method as explained in §5.1.

Six choice questions were required to assess each switching probability. Table 2 gives the procedure followed by the computer program to determine probability  $p_1$ . The process is similar than for the outcome  $x_1$  except that the number of choice questions is 6 instead of 5. Only integers were admitted as feasible midpoint values. In other words, when a midpoint value of an interval is not an integer, the computer program screens/saves the closest integer on the left of this value. Thus, in the third question, the probability 87% was taken as a midpoint value instead of 87.5%. Finally, note that after six questions, the switching probability ends up in the interval  $[85, 87]$ . In this case, the computer program sets  $p_1=86\%$ .<sup>9</sup>

<sup>9</sup> The procedure followed by the computer program to determine  $p_1$  for losses is similar to that described in Table 2, except that the choice pattern AABAAA must be replaced by the choice pattern BBABBB.

Contrary to the assessment procedure for outcomes in TO-experiments, the choice questions allowing the assessment of a given probability  $p$  were not successively sequenced. In fact, to determine probabilities  $p_1, \dots, p_5$  and  $p'_2$  the subject was first asked to successively choose between the prospects  $A_i^1=(x_i, 1)$  and  $B_i^1=(x_6, 1/2; x_0)$ ,  $i=1, \dots, 5$ , and then between the prospects  $A_*^1=(x_3, 1)$  and  $B_*^1=(x_4, 1/2; x_2)$ . The next step consisted of six new questions taking into account the choices expressed in the first step. Thus, if the subject's choices in the first step were, for instance,  $A_1^1, A_2^1, A_3^1, A_4^1, A_5^1$ , and  $A_*^1$ , she was then asked to successively choose between the prospects  $A_i^2=(x_i, 1)$  and  $B_i^2=(x_6, 3/4; x_0)$ ,  $i=1, \dots, 5$ , and also between the prospects  $A_*^2=(x_3, 1)$  and  $B_*^2=(x_4, 3/4; x_2)$ . Overall, each subject was confronted with six series of six choice questions.

Note that with this sequencing of choice questions, the subject can hardly infer that she is in fact asked to match probability  $p$ . This special sequencing of choice questions was designed to avoid the problem of "framed probabilities" (Hershey and Schoemaker 1985) as well as possible biases due to scale compatibility (Tversky et al. 1988, Delqu   1993).

Once the probabilities  $p_1, \dots, p_5, p'_2$  were assessed, the computer program displayed successively the pairs of prospects  $A_i^4, B_i^4, i=1, \dots, 5, A_*^4, B_*^4$ , and asked the subject to designate the preferred prospect in each pair. The aim of this step was to check the subjects' reliability in PW-experiments.

## 5. Results

### 5.1. Reliability

We can test the reliability of subjects' responses by seeing how often they expressed the same preference for the same prospect. As explained above, after each standard sequence assessment every subject was confronted, once again, with the pair of prospects previously presented in the fourth iteration of the bisection processes concerning outcomes and probabilities.

Overall, 19% of the subjects' responses (i.e., 182 out of 960 responses) reveal reversals in preferences. This number is lower than those obtained by

Camerer (1989) and Starmer and Sugden (1989). For TO-experiments, the number of reversals is more important for gains (17.9%) than for losses (13.7%). For PW-experiments, 19.1% and 25% of subjects' responses reveal reversals for gains and losses respectively.<sup>10</sup>

Because responses are chained in the trade-off method (the elicitation of  $x_i$  requires the value  $x_{i-1}$  as input), one may expect errors to propagate similarly as in other chained methods such as the bisection version of the CE method (i.e., fractile method). For the impact of error propagation (on the assessment of  $w$ ) to be small, the subject should be "uniformly consistent" at each stage ( $i=1, \dots, 6$ ) of the assessment of the standard sequence of outcomes. An inconsistency that happens early in the assessment process (i.e., for a small  $i$ ) can induce error propagation in the standard sequence, and therefore, distort the measurement of the probability weighting function. Consequently, a reliable measurement of  $w$  needs both overall consistency and uniform consistency.

The data show that in addition to a rather low overall rate of reversals, subjects consistency did not depend on the rank of the outcome or the probability the experimenter was determining. Indeed, Cochran tests (one-factor ANOVA test with repeated measures for dichotomous data) reveal that reliability of subjects does not change significantly during trade-off processes for gains ( $\chi^2_5 = 3.51 < \chi^2_{5, 0.95} = 11.1$ ), nor for losses ( $\chi^2_5 = 2.68 < 11.1$ ). Similarly, reliability does not change significantly during PW-experiments for gains ( $\chi^2_5 = 5.25 < 11.1$ ) and for losses ( $\chi^2_5 = 7.96 < 11.1$ ).<sup>11</sup> This significant stability in TO-experiments and in PW-experiments is consistent with the idea

that choice-based methods produce "reliable" measurements (e.g., Bostic et al. 1990).

Another possibility to check whether the trade-off method produces genuine standard sequences consists of seeing if  $x_0, x_3, x_6$  and  $x_2, x_3, x_4$  are standard sequences. For that purpose, probabilities  $r$  and  $s$  satisfying  $x_3 \sim (x_6, r; x_0)$  and  $x_3 \sim (x_4, s; x_2)$  were compared for each subject. Clearly, under CPT and the absence of error propagation, we must have  $r=s$ . A paired  $t$  test confirms this equality for gains ( $p=0.51$ ). The mean and standard deviation of the absolute errors (i.e., the  $|r-s|$ 's) are 0.11 and 0.08 respectively. However, for losses, this equality is rejected ( $p=0.01$ ). Here, the mean and standard deviation of the absolute errors are 0.15 and 0.12 respectively. This inconsistency might be explained by the sensitivity of subjects to the variance of losses.<sup>12</sup> Indeed, it will be shown that the elicited standard sequences for losses are quite similar to those recently obtained by Fennema and van Assen (1998). Furthermore, the subjects' consistency was particularly high in TO-experiments for losses.

## 5.2. Utility Functions

One of the most distinctive characteristics of PT and CPT as compared to EU theory is the way the "transformation of outcomes" is taken into account. Indeed, Kahneman and Tversky (1979) assume that the utility function is concave for gains and convex for losses. This assumption is a mere corollary of the psychological principle of *diminishing sensitivity*: The impact of a change diminishes with the distance from the reference point.

For a standard sequence of outcomes  $x_0, \dots, x_6$ , with  $\Delta'_i = |x_i - x_{i-1}|$ ,  $i=1, \dots, 6$ , and  $\Delta''_j = \Delta'_{j+1} - \Delta'_j$ ,  $j=1, \dots, 5$ , the shape of the corresponding utility function depends on the sign of  $\Delta''_j$  when  $j$  varies. Thus, when the above standard sequence of outcomes represents gains (respectively losses), the utility function is concave (respectively convex) if and only if  $\Delta''_j$  is positive (respectively negative) for  $j=1, \dots, 5$ .

<sup>12</sup> Note that the variance of the prospect  $(x_6, r; x_0)$  is higher than the variance of the prospect  $(x_4, s; x_2)$ .

<sup>10</sup> Starmer and Sugden (1989) and Camerer (1989) found respectively that 31.6% and 26.5% of the subjects reversed preference.

<sup>11</sup> These results are consistent with another test of reliability performed in the pilot study for 15 (out of 18) subjects. This test consisted of the assessment of two standard sequences of gains:  $x_0, \dots, x_4$  ( $x_0=1,000$ ,  $r=0$ ,  $R=500$ ,  $p=2/3$ ) and  $y_0, \dots, y_4$  ( $y_0=1,000$ ,  $r=0$ ,  $R=600$ ,  $p=1/2$ ) and two probabilities  $p_x$  and  $p_y$  such that  $x_2 \sim (x_0, p_x; x_4)$  and  $y_2 \sim (y_0, p_y; y_4)$  respectively. A Wilcoxon test shows that there was no significant difference between  $p_x$  and  $p_y$  ( $p < 0.01$ ). The mean and standard deviation of the absolute errors (i.e., the  $|p_x - p_y|$ 's) were 0.095 and 0.083 respectively.

Table 3 Classification of Subjects

		LOSSES				Total
		Concave	Convex	Linear	Mixed	
GAINS	Concave	4	13	4	0	21
	Convex	1	0	3	4	8
	Linear	3	4	0	0	7
	Mixed	0	0	3	1	4
	Total	8	17	10	5	40

Recent experimental studies on elicitation of utility functions for monetary outcomes (e.g., Wakker and Deneffe 1996) show that the trade-off method needs high outcomes to obtain utility functions with pronounced curvature. The outcomes used in the experimental study reported in this paper are lower than those used by the authors of the above mentioned paper. For less pronounced curvature, the sign of the  $\Delta_j''$ 's is expected to be less "pronounced," and therefore, very sensitive to response error.

Because of response error, subjects with three (out of five) or more positive (respectively negative)  $\Delta_j''$ 's were classified as exhibiting a concave (respectively convex) utility function for gains (respectively losses).<sup>13</sup> Subjects were similarly classified as exhibiting a convex or linear utility function. When a utility function could not be classified by the means of the above criteria, it was considered as exhibiting a mixed shape. Table 3 gives the subjects classification for gains and losses.<sup>14</sup>

The results for gains are not in contradiction with the principle of diminishing sensitivity. Table 3 shows that 21 out of 40 elicited utility functions exhibit a concave shape, whereas 8 convex and 7 linear utility functions were obtained for the remainder of subjects.<sup>15</sup> Four elicited utility functions exhibit a mixed

Table 4 Averaged Results for Gains and Losses ( $n=40$ )

$i$	GAINS		LOSSES	
	$x_i$	$\Delta_i'$	$ x_i $	$\Delta_i'$
1	3190 (1158)	2190	2027 (667)	1027
2	5465 (2123)	2275	3162 (1437)	1135
3	7945 (3066)	2480	4225 (1991)	1062
4	10512 (4059)	2567	5550 (2751)	1325
5	13002 (5239)	2490	6872 (3648)	1322
6	15680 (6221)	2677	8240 (4594)	1367

Standard deviations are between parentheses.

shape. When only concave and convex shapes are taken into account, we can conclude that there are significantly more concave utility functions than convex ones ( $p=0.01$ , binomial, one-tailed). Averaged results for gains are given in Table 4.

For losses, 17 subjects out of 40 have convex utility functions, whereas 8 subjects have concave utility functions. Ten subjects were classified as having linear utility functions.<sup>16</sup> The remainder of the subjects show mixed shapes. The number of convex cases is, however, significantly higher than the number of concave ones ( $p=0.025$ , binomial, one-tailed). Table 4 shows averaged results for losses.

When only subjects having nonlinear utility functions for gains and for losses are taken into account, only 13 out of 40 subjects exhibit concavity for gains and convexity for losses. This low rate of consistency with the hypothesis of diminishing sensitivity may be explained, in part, by the relatively large number of linear cases (4 for gains and 4 for losses, Table 3). Furthermore, the tendency of the elicited utility functions to exhibit a linear shape might decline if the amounts of money are sufficiently important to allow diminishing sensitivity to both gains and losses.

When a power utility function is assumed (Tversky and Kahneman 1992), median estimates of  $\alpha$  and  $\beta$  (using standard nonlinear least squares regression) are 0.89 and 0.92 respectively. These estimates are consistent with the hypothesis of diminishing sensitivity.

<sup>13</sup> Fennema and van Assen (1998) and Bleichrodt and Pinto (1998) use similar criteria.

<sup>14</sup> Due to the fact that each  $x_i$  was chosen in an interval (as shown in Table 2), a  $\Delta_j''$  was considered as positive (respectively negative) when  $\Delta_j'' \geq 100$  (respectively  $\Delta_j'' \leq -100$ ).  $\Delta_j''$  was given a zero value when  $-100 < \Delta_j'' < 100$ .

<sup>15</sup> There are 21 (14) subjects with at least 3 (4) positive values of  $\Delta_j''$ , 8 (5) subjects with at least 3 (4) negative values of  $\Delta_j''$ , and 7 (5) subjects with at least 3 (4) zero values of  $\Delta_j''$ .

<sup>16</sup> There are 17 (12) subjects with at least 3 (4) negative values of  $\Delta_j''$ , 8 (4) subjects with at least 3 (4) positive values of  $\Delta_j''$ , and 10 (7) subjects with at least 3 (4) zero values of  $\Delta_j''$ .

They are close to those obtained by Tversky and Kahneman (1992).<sup>17</sup>

### 5.3. Probability Weighting Functions

For Tversky and Kahneman (1992), the most distinctive characteristic of individual behavior under risk is what they called the *fourfold pattern* of risk attitude. It claims that individuals exhibit: (i) risk seeking for gains and risk aversion for losses of low probability; (ii) risk aversion for gains and risk seeking for losses of high probability. The authors remark that since the fourfold pattern arises over a wide range of outcomes, it cannot be explained by the utility of money as suggested by Friedman and Savage (1948) and Markowitz (1952). Instead, they infer that this fourfold pattern of risk attitude suggests an inverse S-shaped nonlinear transformation of the probability scale as well as an S-shaped utility function. Tversky and Kahneman (1992) recognize, however, that instead of clarity of the overall pattern, the individual data reveal both noise and individual differences.

Probability weighting functions were elicited for 40 subjects both for gains and losses. The assessed values were inverse images  $(w)^{-1}(p)$ , rather than images  $w(p)$  for  $p=i/6$ ,  $i=1, \dots, 5$  (see Figure 4). Table 5 gives medians, means, and standard deviations of the corresponding empirical distributions. Note that medians and means are higher for gains than for losses. This suggests that probability weighting functions curves for losses might lie significantly above those of the probability weighting functions for gains.

One-tailed sign tests of  $H_0: w(p)=p$  in Table 6 show that subjects indeed transform probabilities both for gains and for losses.<sup>18</sup> More precisely, overweighting of small probabilities and underweighting of moderate and high probabilities are confirmed by data except for  $p=2/6$  for both gains and losses and for  $p=3/6$  for losses.

As noted in §3.2, LS can be tested straightforwardly through the inequalities:  $\Delta_{w^{-1}}(1) < \Delta_{w^{-1}}(2)$  and  $\Delta_{w^{-1}}(1) < \Delta_{w^{-1}}(3)$  where  $\Delta_{w^{-1}}(i) = w^{-1}(\frac{i}{6}) - w^{-1}(\frac{i-1}{6})$ . One-tailed paired  $t$  tests confirm the presence of LS for gains and losses (Table 7, Rows 1,2). Furthermore, the elicited weighting functions exhibit US. Alternative hypotheses  $\Delta_{w^{-1}}(6) < \Delta_{w^{-1}}(5)$  and  $\Delta_{w^{-1}}(6) < \Delta_{w^{-1}}(3)$  are confirmed for gains and for losses (Table 7, Rows 3,4). In addition, the alternative hypothesis  $\Delta_{w^{-1}}(6) < \Delta_{w^{-1}}(1)$  is confirmed by the observations for gains and for losses (Table 7, Row 5). Consequently, the transformation of probabilities has more impact near certainty than near impossibility. This finding is consistent with Tversky and Fox (1995).

To sum up, the elicited weighting functions are consistent with the fourfold pattern. They exhibit overweighting of small probabilities and underweighting of moderate and high probabilities along with LS and US.

One advantage of the elicitation of weighting functions for gains and losses is the possibility of comparing the two functions without passing through a parametric form which may influence the result of the comparison. The elicited  $w^+$  and  $w^-$  have similar shapes, but they are not identical for gains and losses. Table 8 shows that the hypothesis of identity of weighting functions for gains and losses, called *reflection*, is not rejected near  $p=0$  and  $p=1$ , and clearly rejected for moderate probabilities. Furthermore, one-tailed paired  $t$  tests show that the weighting function for losses is significantly "above" the weighting function for gains for moderate probabilities at the 0.025 level. In other words,  $w^-$  exhibits more *elevation* than  $w^+$ . Table 8 shows also that  $(w^+)^{-1}(p)$  and  $(w^-)^{-1}(p)$  are more correlated on the first half of the unit interval than on its second half.

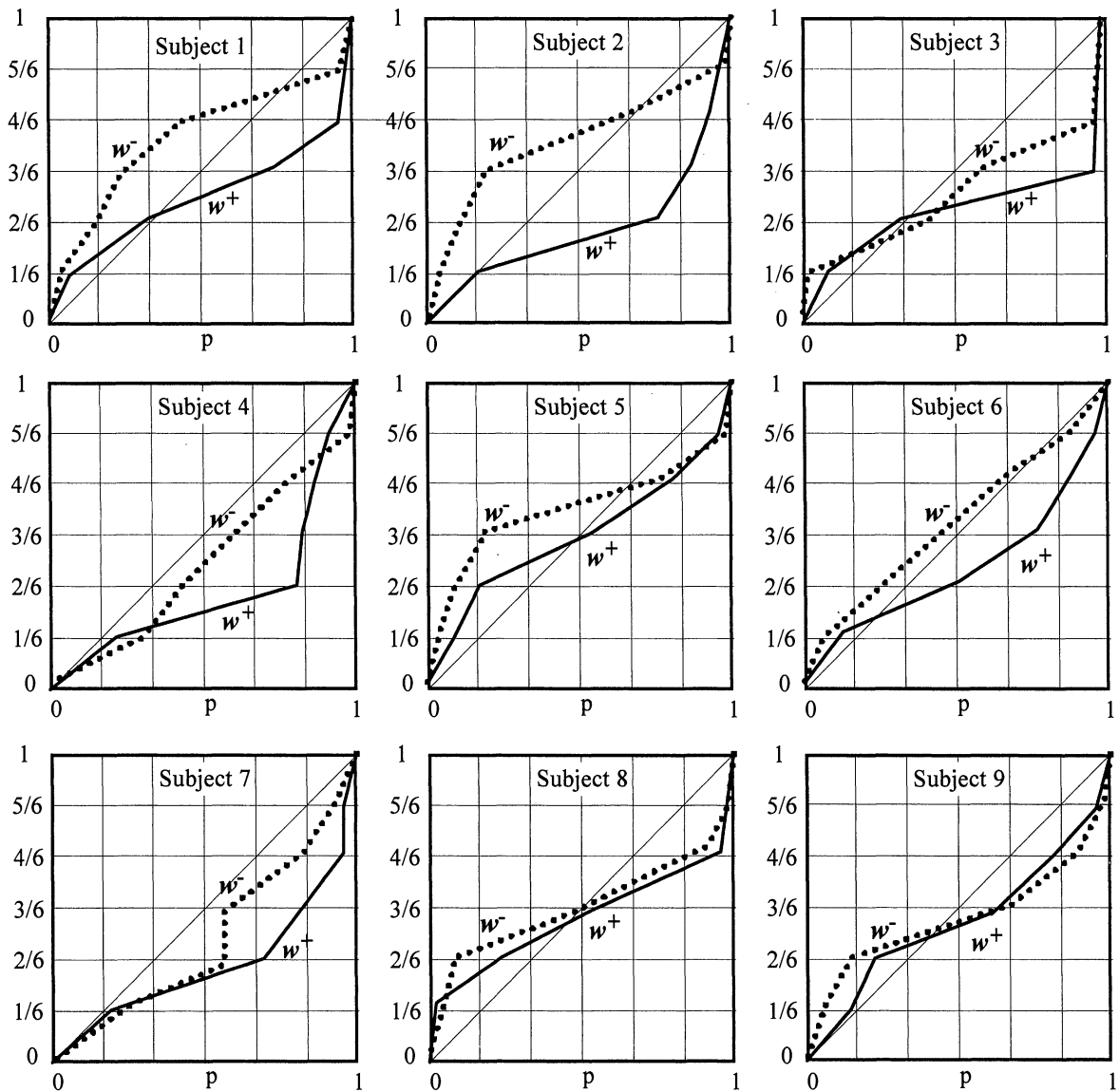
Camerer (1992) and Abdellaoui and Munier (1998) obtained results favorable to EU and "locally" favorable to EU respectively, inside the unit triangle. In fact, accepting EU inside the unit triangle implies a linear shape of the weighting function over a subset of the interval of middle probabilities. Nevertheless, Wu and Gonzalez (1996) report (parametric and nonparametric) results indicating that the weighting function presents significant curvature strictly within the boundaries of the unit interval. Their data-fitting

<sup>17</sup> For gains, Wu and Gonzalez (1996) obtained  $\alpha=0.52$ .

<sup>18</sup> Because probability  $p=1/6$  was not feasible in the assessment process ( $p=0.166\dots$ ), equality  $w(p)-p=0$  is considered as satisfied when  $p$  belongs to the interval  $[16/100, 17/100]$ . A similar convention is used for  $p=2/6, 4/6, 5/6$ .



Figure 4 Elicited Weighting Functions for the First Nine Subjects



results indicate that a nonlinear weighting function within the boundaries of the unit interval outperforms a linear weighting function with discontinuities at 0 and 1.

The hypothesis of linearity of the probability weighting function away from the boundaries of the unit interval, consistent with  $\Delta_{w^{-1}}(k) = w^{-1}(k/6) - w^{-1}((k-1)/6) = \text{constant}$  for  $k=2, 3, 4, 5$ , is clearly rejected by a

Friedman test for gains ( $\chi^2_3 = 16.71, p < 0.001$ ) as well as for losses ( $\chi^2_3 = 9.41, p < 0.05$ ). Note that the rejection is "more pronounced" for gains than for losses. A similar test of the hypothesis  $\Delta_{w^{-1}}(k) = \text{constant}$  over the interval  $[1/2, w^{-1}(5/6)]$ , i.e.,  $k=3, 4, 5$ , confirm the nonlinearity of the probability weighting function for gains ( $\chi^2_2 = 16.28, p < 0.001$ ) and for losses ( $\chi^2_2 = 6.78, p < 0.05$ ). Due to the flatness of  $w$  over  $[w^{-1}(1/6), 1/2]$ , linearity



**Table 5** Medians, Means and Standard Deviations for  $(w^+)^{-1}(p)$  and  $(w^-)^{-1}(p)$

Probability $p$	$(w^+)^{-1}(p)$			$(w^-)^{-1}(p)$		
	Median	Mean	S.D.	Median	Mean	S.D.
1/6	0.150	0.160	0.128	0.133	0.125	0.095
2/6	0.350	0.409	0.229	0.300	0.293	0.161
3/6	0.650	0.661	0.175	0.530	0.502	0.206
4/6	0.895	0.849	0.108	0.782	0.781	0.160
5/6	0.955	0.947	0.048	0.955	0.934	0.062

**Table 6** Counts of  $w(p) - p > 0$  and  $w(p) - p < 0$  (One-Tailed Sign Tests)

$w(p) - p$	GAINS		LOSSES	
	$> 0$	$< 0$	$> 0$	$< 0$
$p = 1/6$	24*	10	26*	12
$p = 2/6$	17 <sup>ns</sup>	23	19 <sup>ns</sup>	20
$p = 3/6$	10	27**	24 <sup>ns</sup>	15
$p = 4/6$	4	33**	9	28**
$p = 5/6$	1	35**	2	36**

ns: nonsignificant for  $\alpha = 0.05$ ; \* $p < 0.05$ ; \*\* $p < 0.01$ .

**Table 7** Lower and Upper Subadditivity (One-Tailed Paired  $t$  Tests)

Alternative hypothesis	GAINS	LOSSES
	$t_{39}$	$t_{39}$
$\Delta_{w^{-1}}(1) < \Delta_{w^{-1}}(2)$	-3.06**	-2.33*
$\Delta_{w^{-1}}(1) < \Delta_{w^{-1}}(3)$	-2.22*	-2.81**
$\Delta_{w^{-1}}(6) < \Delta_{w^{-1}}(5)$	-4.20***	-4.53***
$\Delta_{w^{-1}}(6) < \Delta_{w^{-1}}(3)$	-6.67***	-4.84***
$\Delta_{w^{-1}}(6) < \Delta_{w^{-1}}(1)$	-4.81***	-3.35**

\* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ .

can be tested only for  $k=2, 3$ . A paired  $t$  test shows that the hypothesis  $\Delta_{w^{-1}}(2) = \Delta_{w^{-1}}(3)$  is consistent with data for gains ( $t_{39} = -0.06$ ,  $p = 0.9$ ) and for losses ( $t_{39} = -1.03$ ,  $p = 0.2$ ). It must be recognized, however, that a more efficient curvature detection on the first half of the unit interval needs the assessment of more points ( $p, w^{-1}(p)$ ).

In their attempt to give a psychological interpretation to the weighting function, Gonzalez and Wu (1999) distinguish two logically independent psychological properties. Discriminability refers to how people discriminate intermediate probabilities. It is indexed by the *curvature* of  $w$ .

**Table 8** Two-Tailed Paired  $t$  test and Correlations

$(w^+)^{-1}(p)$ vs $(w^-)^{-1}(p)$	$t_{39}$	Correlation
$p = 1/6$	1.7 <sup>ns</sup>	0.35
$p = 2/6$	3.19**	0.35
$p = 3/6$	4.81***	0.40
$p = 4/6$	2.38*	0.11
$p = 5/6$	1.12 <sup>ns</sup>	0.09

ns: nonsignificant for  $\alpha = 0.05$ .

\* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ .

**Table 9** Parametric Estimates for Median Data

	Parametric specifications:		
	$\frac{p^\gamma}{[p^\gamma + (1-p)^\gamma]^{\frac{1}{\gamma}}}$	$\frac{\delta p^\gamma}{\delta p^\gamma + (1-p)^\gamma}$	$\gamma$
GAINS	0.60 (0.027)	0.65 (0.054)	0.60 (0.050)
LOSSES	0.70 (0.028)	0.84 (0.054)	0.65 (0.049)

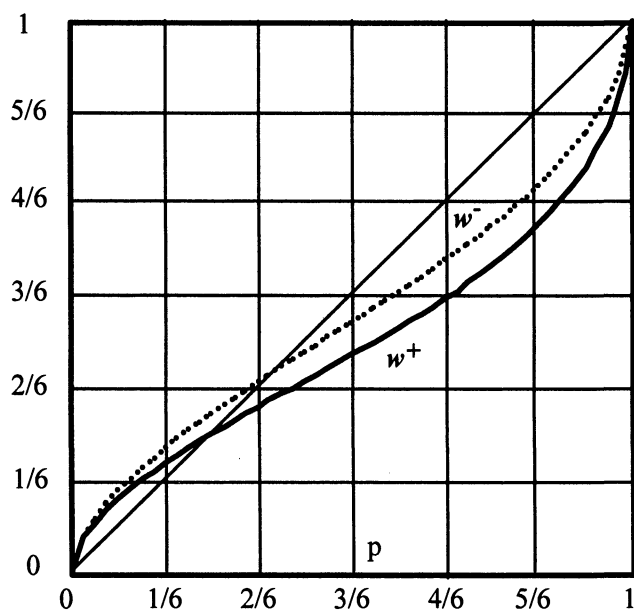
Standard Errors are in parentheses.

Attractiveness refers to the degree of over/under-weighting; it is indexed by the *elevation* of  $w$ . The authors note that the Lattimore et al. (1992) function takes into account curvature and elevation through its two parameters  $\gamma$  and  $\delta$  respectively.<sup>19</sup>

Table 9 gives parameter estimates of the probability weighting functions for median data (using standard nonlinear least squares with a normally distributed error term). When the weighting function is assumed to have a Tversky and Kahneman's single-parameter form, the median estimates obtained are very close to those obtained by Tversky and Kahneman (1992) for gains and losses, which were 0.61 and 0.69 respectively. Nevertheless, the Lattimore et al. (1992) function provides a clearer separation between losses and gains. The estimates of  $\delta$  show that the probability weighting function for median data exhibit more elevation for losses than for gains (see Figure 5). Such a clear distinction is not allowed by Tversky and Kahneman's one-parameter specification of  $w$ . Note that if  $(\delta^+, \gamma^+)$  and  $(\delta^-, \gamma^-)$  designate the parameters of  $w^+$  and  $w^-$  respectively, duality (i.e.,  $w^+(p) = 1 - w^-(1-p)$ )

<sup>19</sup> Kilka and Weber (1998) extend these concepts to uncertainty.

Figure 5 Weighting Functions Estimates Based on Median Data



implies that  $\gamma^+ = \gamma^-$  and  $\delta^+ = 1/\delta^-$ <sup>20</sup>. Parameter estimates for median data do not meet these conditions.

At the level of individual subjects, elevation is also more pronounced for losses than for gains (i.e.,  $\delta$  for losses is greater than  $\delta$  for gains). This is true in the present experimental study for 27 subjects out of 40 ( $p < 0.05$ , one-tailed sign test). Figure 6 shows the distributions of the estimates of  $\delta$  and  $\gamma$  respectively.

## 6. Discussion and Conclusion

Economists assume diminishing marginal utility, and therefore, concavity of utility functions. Psychologists, on the other hand, consider that due to the principle of diminishing sensitivity, the utility function is concave for gains and convex for losses. The experimental results reported in this paper support the hypothesis of diminishing sensitivity for gains and losses. They are very similar to those obtained by Wakker and Deneffe (1996) and Fennema and van Assen (1998) respectively. Also, it is remarkable that the experimental findings

for utility functions using gambles with specified probabilities are similar to those obtained by Wakker and Deneffe (1996) in a context and ambiguity (i.e., without probabilities).

The main result for weighting functions is that the fundamental hypotheses of LS and US are not rejected. This "nonparametric" confirmation of subadditivity agrees with the results obtained through parametric methods in previous experimental studies.<sup>21</sup> It is also a direct observation of the phenomenon of nonlinearity in probabilities of preferences under risk. Indeed, as predicted by the psychological principle of diminishing sensitivity, the two end points of the probability interval serve as reference points. Therefore, increments near these points of the probability interval have more impact than increments in the middle of this interval. Furthermore, the data show that the transformation of probabilities has more impact near certainty than near impossibility.

Away from the end points of the unit interval, linearity of the probability weighting function is rejected. More specifically, the data reported in this paper locates the main source of nonlinearity in the second half (of the interior) of the unit interval. It must be recognized, however, that a more efficient testing of linearity of  $w$  needs the assessment of more points ( $p, w^{-1}(p)$ ) than has been done in the present paper.

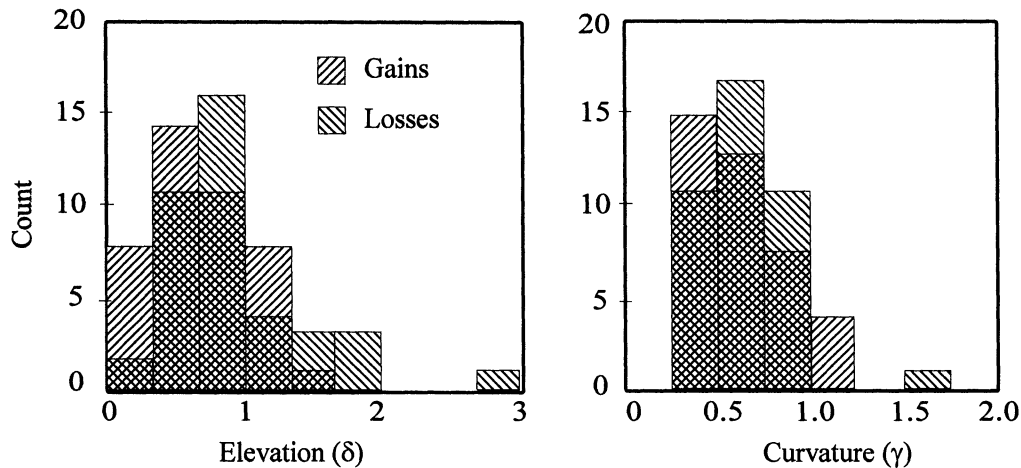
The elicited weighting function for gains and the elicited weighting function for losses are systematically different. This confirms some early empirical findings showing that individuals distort probabilities differently when gains are transformed into losses in risky choice situations (e.g., Cohen et al. 1987, Currim and Sarin 1989, Abdellaoui 1995b). In fact, weighting functions for losses exhibit significantly more elevation than weighting functions for gains.

The empirical results obtained in this paper also have important implications for RDEU and EU. The introduction of two weighting functions in CPT seems more appropriate than duality (assumed in RDEU) for taking into account the observed tendency of individuals to treat probabilities differently when

<sup>20</sup> This follows from the equality:

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

<sup>21</sup> Bleichrodt and Pinto (1998) and Gonzalez and Wu (1999) obtained similar results.

Figure 6 Distributions of  $\delta$  and  $\gamma$  for Gains and Losses

passing from gains to losses and vice versa. These experimental findings can also be seen as a confirmation that the traditional utility elicitation methods can produce "distorted" utility functions for expected utility prescriptive purposes. Indeed, the PE method, the CE method, as well as the lottery equivalent method suggested by McCord and de Neufville (1986), assume EU, and therefore, are subject to the disturbing effects of probability distortion. The trade-off method has the main advantage of avoiding these disturbing effects. Its main disadvantage in theory is that it allows error propagation.

The two-step procedure described and used in this paper has the advantage of simplicity. It needs only a few choices to elicit both the utility and the probability weighting function. Furthermore, the probability weighting function is obtained through a simple formula inferred from an indifference between a sure outcome and a binary lottery. It must be recognized, however, that in theory error propagation in the trade-off method can produce "noisy" probability weighting functions. Subsection 5.1 gives some reassuring results, particularly concerning error propagation in the trade-off method, although the mean values of  $|r - s|$  seem somewhat large (about 0.10 for gains).

A better performance of this two-step elicitation method regarding error propagation can be expected through individual interviews using more cross-

checking to reduce the impact of errors. Furthermore, an appropriate assessment of more points ( $p, w^{-1}(p)$ ) than done in the present paper can also be used to systematically investigate individual differences in probability weighting. Individual differences can hardly be explained by "noise". This is confirmed in Gonzalez and Wu (1999). These authors report interpersonal comparisons based on discriminability and attractiveness for 10 subjects, showing a considerable variation on these two dimensions.<sup>22</sup>

<sup>22</sup> This paper received helpful comments from Carolina Barrios, Han Bleichrodt, Philippe Delquié, Jean-Yves Jaffray, Duncan Luce, Bertrand Munier, Jason Shachat, Marcel van Assen, Peter Wakker and George Wu. I also thank anonymous referees for especially useful and detailed comments. Support by a grant from the Ecole Normale Supérieure de Cachan (Department of Economics and Management) is gratefully acknowledged.

## References

- Abdellaoui, M. 1995a. Estimating probability transformation functions under RDEU. Unpublished manuscript, GRID, ENS de Cachan, France.
- . 1995b. Comportements individuels devant le risque et transformation des probabilités. *Revue d'Economie Politique* 105(1) 157-178.
- . 1996. Is the probability transformation function in RDEU models sensitive to the magnitude of losses? Working paper, GRID, ENS de Cachan, France.
- , B. Munier 1998. The risk-structure dependence effect: Experimenting with an eye to decision-aiding. *Ann. Oper. Res.* 80 237-252.

- Allais, M. 1988. The general theory of random choices in relation to the invariant cardinal utility function and the specific probability function. B. Munier, ed. *Risk and Rationality*. Reidel, Dordrecht, 233–289.
- Bleichrodt, H., J. L. Pinto. 1998. A parameter-free elicitation of the probability weighting function in medical decision analysis. Working paper, Erasmus University, The Netherlands.
- Bostic, R., R. J. Herrnstein, R. D. Luce. 1990. The effect on the preference reversal phenomenon of using choice indifference. *J. Econom. Behavior and Organ.* 13 193–212.
- Camerer, C. F. 1989. An experimental test of several generalized utility theories. *J. Risk and Uncertainty* 2 61–104.
- . 1992. Recent tests of generalizations of expected utility theory. W. Edwards, ed. *Utility: Theories, Measurement and Applications*. Kluwer Academic Publishers, Boston, MA.
- , T.-H. Ho. 1994. Nonlinear weighting of probabilities and violations of the betweenness axiom. *J. Risk and Uncertainty* 8 167–196.
- Cohen, M., J. Y. Jaffray. 1988. Preponderance of the certainty effect over probability distortion in decision making under risk. B. R. Munier, ed. *Risk and Rationality*. Reidel, Dordrecht, 173–187.
- , —, T. Said. 1987. Experimental comparisons of individual behavior under risk and under uncertainty for gains and for losses. *Organ. Behavior and Human Decision Process.* 39 1–22.
- Currim, I., R. K. Sarin. 1989. Prospect versus utility. *Management Sci.* 35 22–41.
- Delquié, P. 1993. Inconsistent trade-offs between attributes: New evidence in preference assessment biases. *Management Sci.* 39 1382–1395.
- Fennema, H., M. van Assen. 1998. Measuring the utility of losses by means of the trade-off method. *J. Risk and Uncertainty* 17 277–295.
- Fox, C., A. Tversky. 1998. A belief-based account of decision under uncertainty. *Management Sci.* 44 897–895.
- Friedman, M., L. J. Savage. 1948. The utility analysis of choices involving risk. *Political Econom.* 56(4) 279–304.
- Goldstein, W. M., H. J. Einhorn. 1987. Expression theory and the preference reversal phenomena. *Psych. Rev.* 94 236–254.
- Gonzalez, R., G. Wu. 1999. On the shape of the probability weighting function. *Cognitive Psych.* 38 129–166.
- Harless, D. W., C. Camerer. 1994. The predictive utility of generalized expected utility theories. *Econometrica* 62 1251–1289.
- Hershey, J. C., P. H. J. Schoemaker. 1985. Probability versus certainty equivalence methods in utility measurement: Are they equivalent? *Management Sci.* 31 1213–1231.
- Hey, J. D., C. Orme. 1994. Investigating generalizations of expected utility theory using experimental data. *Econometrica* 62 1296–1326.
- Kahneman, D., A. Tversky. 1979. Prospect theory: An analysis of decision under risk. *Econometrica* 47 263–291.
- Kilka, M., Weber, M. 1998. What determines the shape of the probability weighting function under uncertainty. Working paper, Universität Mannheim, Germany.
- Lattimore, P. M., J. R. Baker, A. D. Witte. 1992. The influence of probability on risky choice. *J. Econom. Behavior and Organ.* 17 377–400.
- Markowitz, H. 1952. The utility of wealth. *J. Political Econom.* 60 151–158.
- McCord, M., R. de Neufville. 1986. Lottery equivalents: Reduction of the certainty effect problem in utility assessment. *Management Sci.* 32 56–60.
- Miyamoto, J. M. 1988. Generic utility theory: Measurement foundations and applications in multiattribute utility theory. *J. Math. Psych.* 32 357–404.
- Prelec, D. 1998. The probability weighting function. *Econometrica* 66 497–527.
- Quiggin, J. 1982. A theory of anticipated utility. *J. Econom. Behavior and Organ.* 3 323–343.
- Segal, U. 1987. Some remarks on Quiggin's anticipated utility. *J. Econom. Behavior and Organ.* 8 145–154.
- Starmer, C., R. Sugden. 1989. Violations of the independence axiom in common ratio problems: An experimental test of some competing hypotheses. *Ann. Oper. Res.* 19 79–102.
- Thaler, R. H., E. Johnson. 1990. Gambling with the house money and trying to break even: The effects of prior outcomes on risky choice. *Management Sci.* 36 643–660.
- Tversky, A., C. Fox. 1995. Weighting risk and uncertainty. *Psych. Rev.* 102 269–283.
- , D. Kahneman. 1992. Advances in prospect theory: Cumulative representation of uncertainty. *J. Risk and Uncertainty* 5 297–323.
- , P. P. Wakker. 1995. Risk attitudes and decision weights. *Econometrica* 63(6) 297–323.
- , S. Sattah, P. Slovic. 1988. Contingent weighting in judgement and choice. *Psych. Rev.* 95 371–384.
- Wakker, P. P. 1989. *Additive Representations of Preferences: A New Foundation in Decision Analysis*. Kluwer Academic Publishers, Dordrecht.
- . 1994. Separating marginal utility and probabilistic risk aversion. *Theory and Decision* 36 1–44.
- . 1997. Testing and characterizing properties of nonadditive measures. Mimeo., University of Tilburg, The Netherlands.
- , D. Deneffe. 1996. Eliciting von Neumann-Morgenstern utilities when probabilities are distorted or unknown. *Management Sci.* 42 1131–1150.
- Wu, G., R. Gonzalez. 1996. Curvature of the probability weighting function. *Management Sci.* 42 1676–1690.
- , —. 1998. Common consequence conditions in decision making under risk. *J. Risk and Uncertainty* 16 115–139.

Accepted by Martin Weber; received March 15, 1998. This paper was with the author 5 months for 2 revisions.