# Choice-based Elicitation and Decomposition of Decision Weights for Gains and Losses under Uncertainty

Mohammed Abdellaoui[a], Frank Vossmann[b], Martin Weber[c]

February 4, 2004 (Revised)

**Abstract:**

This paper reports the results of an experimental parameter-free elicitation and decomposition of decision weights under uncertainty. Assuming cumulative prospect theory, utility functions were elicited for gains and losses at an individual level using the trade-off method. Subsequently, decision weights were elicited through certainty equivalents of uncertain two-outcome prospects. Furthermore, decision weights were decomposed using observable choice instead of invoking other empirical primitives, as in previous experimental studies. The choice-based elicitation of decision weights allows for a quantitative study of their characteristics, and also allows, among other things, for the examination of the sign-dependence hypothesis for observed choice under uncertainty. Our results confirm concavity of the utility function in the gain domain and bounded subadditivity of decision weights and choice-based subjective probabilities. We also find evidence for sign dependence of decision weights.

**Keywords:**

*Decision under Uncertainty, Choquet Expected Utility, Cumulative Prospect Theory, Decision Weights, Choice-based Probabilities, Probability Weighting.*

---

[a] GRID-CNRS, Ecole Nationale Supérieure d'Arts et Métiers, Maison de la Recherche de l'ESTP, 30 Avenue du Président Wilson, 94230 Cachan – France, abdellaoui@grid.ensam.estp.fr; [b] Lehrstuhl für Bankbetriebslehre, Universität Mannheim, 68131 Mannheim – Germany, vossmann@bank.bwl.uni-mannheim.de; [c] Lehrstuhl für Bankbetriebslehre, Universität Mannheim, 68131 Mannheim – Germany, weber@bank.bwl.uni-mannheim.de.

We thank Chris Starmer, Peter Wakker, and seminar participants at the 8th BDRM Conference in Chicago and at the 2002 European ESA Meeting in Strasbourg for helpful comments. Financial support from the Deutsche Forschungsgemeinschaft (DFG) is gratefully acknowledged.

## 1. Introduction

Subjective expected utility (SEU) theory (Savage 1954) evaluates an uncertain alternative as the sum of utilities from outcomes weighted by the corresponding subjective probabilities, and consequently establishes a simple and intuitive separation of value and belief. If this theory is abandoned in favor of more general theories that replace probabilities by nonadditive decision weights, the above-mentioned separation between value and belief becomes less clear. Among the most influential of these theories are Choquet expected utility (CEU) theory (Gilboa 1987, Schmeidler 1989) and Cumulative Prospect Theory (CPT) (Tversky and Kahneman 1992). CPT is more general than CEU because decision weights are sign-dependent (i.e., depend on whether the associated consequence is a gain or a loss). Under these theories, the decision weight attached to an event depends on the rank-ordering of outcomes in the uncertain alternative under consideration. In other words, it reflects additional considerations attached to the particular decision context, related to decision attitude, over and above pure belief (Wakker 2004).

Researchers suggested decomposing decision weights into a component of belief (e.g., a subjective probability), independent of the particular decision context, and a component reflecting decision attitude. Such a decomposition was initially proposed by Fellner (1961). Subsequently, Tversky and Fox (1995), Fox et al. (1996), Fox and Tversky (1998), Wu and Gonzalez (1999), and Kilka and Weber (2001) used judged probabilities, i.e., degrees of belief resulting from direct judgment and not from choice, to operationalize the decomposition in experimental investigations. More recently, a choice-based decomposition was proposed by Wakker (2004). This approach allows for an elicitation and decomposition of decision weights under uncertainty which is consistent with behavioral concepts used in standard economics.

This paper reports the results of an experimental study in which CPT decision weights are elicited under uncertainty, for gains and losses, at an individual level through a parameter-free elicitation method based on choice questions. Decision weights are then decomposed using choice-based degrees of belief, called choice-based probabilities, as in the recent theoretical decomposition in Wakker (2004). We thus estimate the transformation function that maps choice-based probabilities to decision weights, separately for gains and losses. Below, we compare our approach to those used in recent experimental investigations on decision weights under uncertainty.

Decision weights have been investigated using two research strategies. The first strategy consists of testing simple preference conditions to obtain information about the shape of the weighting function (Wu and Gonzalez 1999). The second strategy consists of eliciting decision weights from individual preferences. Tversky and Fox (1995) and Fox et al. (1996) used a certainty equivalent method and two-outcome prospects to elicit individual decision weights in the gain domain though utility functions were not elicited at the individual level. Indeed, Tversky and Fox (1995) used the same power utility function for all subjects, while Fox et al. (1996) used a similar methodology with a linear utility function (inferred from a preceding experimental study). Subsequently, Wakker and Deneffe (1996) proposed the trade-off method as a utility function elicitation technique which is applicable for known as well as unknown probabilities. With this method it is possible to elicit utilities, at an individual level, without potential distortions due to nonlinear decision weights, and then elicit decision weights in a second step. This methodology of non-parametric elicitation of decision weights can be useful if one desires to systematically test for properties of decision weights without appeal to assumptions regarding the shape of the individual's utility function, its parametric form or the necessity to use more than two-outcome prospects.

To sum up, the present paper extends the approach used by Abdellaoui (2000) and Bleichrodt and Pinto (2000) from risk to uncertainty, and experimentally operationalizes the choice-based decomposition of decision weights proposed by Wakker (2004). Its contribution is complementary to previous experimental investigations (e.g., Fox et al. 1996 and Wu and Gonzalez 1999). For the gain domain, it allows for a verification of the robustness of some results of these experimental studies. This verification is also a means by which we can compare the results of decision weight decompositions based on standard behavioral concepts and those resulting from a decomposition based on non-behavioral concepts studied in psychology. The confirmation of the results previously obtained through the latter approach is fruitful for all fields involved. For the loss domain, our results shed light on a less investigated side of individual decision making under risk and under uncertainty. They particularly allow for a straightforward comparison between the properties of decision weights for gains and losses.

The present paper is structured as follows. Section 2 briefly reviews CPT and introduces the two-stage approach linking decision making under uncertainty to decision making under risk. Section 3 describes the procedure to successively elicit the utility function and decision weights. Section 4 completes the description of our experimental set-up. The results of our study are presented in Section 5. The paper concludes with a summary and discussion of the main findings.

## 2. Theoretical Framework

In decision theory, uncertainty is modeled through a set $S$, called the *state space*. The decision maker knows that exactly one of these states will obtain, but she does not know which one. Subsets of $S$ are called *events*. The objects of choice in this paper are binary *prospects* $(x_1, A; x_2)$ yielding a monetary outcome $x_1$ if event $A$ obtains and a monetary outcome $x_2$ otherwise. Outcomes are expressed as changes with respect to the status quo, i.e. gains or losses. A prospect that involves both a gain and a loss outcome is called *mixed*. Other prospects are called *non-mixed*.

If the prospect $(x_1, A; x_2)$ is evaluated according to CEU, and $x_1 \geq x_2$, its value is given by

(1) $$\pi_1 \cdot u(x_1) + \pi_2 \cdot u(x_2) \ ,$$

where $\pi_1$ and $\pi_2$ are *decision weights* depending on the ranking of outcomes and $u(.)$ is a strictly increasing *utility function* from $\mathbb{R}$ to $\mathbb{R}$. The decision weights $\pi_1$ and $\pi_2$ are defined by

(2) $$\pi_1 = W(A) \ \text{ and } \ \pi_2 = 1 - W(A),$$

where $W(.)$ is a weighting function, i.e. a function from $2^S$ to $[0, 1]$ satisfying monotonicity with respect to set inclusion ($A \subset B \Rightarrow W(A) \leq W(B)$) and the normalization conditions $W(\emptyset) = 0$ and $W(S) = 1$. Under SEU, due to additivity of $W(.)$, decision weights coincide with subjective probabilities.

Under CPT, the utility function satisfies $u(0) = 0$ and decision weights are determined through two weighting functions: $W^+(.)$ for gains and $W^-(.)$ for losses. The value of the prospect $(x_1, A; x_2)$ is still given by Equation (1). If it involves only gains [losses] with $x_1 \geq x_2 \geq 0$ [$x_1 \leq x_2 \leq 0$], $W(.)$ is replaced by $W^+(.)$ [$W^-(.)$] in the decision weights $\pi_1$ and $\pi_2$ defined by Equations (2). For mixed prospects with $x_1 > 0 > x_2$, decision weights are defined by $\pi_1 = W^+(A)$ and $\pi_2 = W^-(S - A)$.

If the utility function satisfies $u(0) = 0$ and the *duality condition* $W^-(A) = 1 - W^+(S - A)$ holds for all events $A$, CPT and CEU coincide. For non-mixed prospects, the model given by Equation (1) agrees with the multiple priors model (Gilboa and Schmeidler 1989) and Gul's (1991) disappointment theory.

Decision under risk can be seen as a special case of decision under uncertainty in which events are replaced by objective probabilities. A prospect ($x_1$, $p$; $x_2$) yields monetary outcome $x_1$ [$x_2$] with probability $p$ [$1 - p$]. In the evaluation of risky prospects, the weighting function $W(.)$ in Equation (2) is replaced by a *probability weighting function* $w(.)$, i.e. a strictly increasing function from [0, 1] to [0, 1] satisfying $w(0) = 0$ and $w(1) = 1$. Analogously, under CPT the weighting functions $W^+(.)$ and $W^-(.)$ are replaced by $w^+(.)$ and $w^-(.)$, respectively.

Following Schmeidler (1989), convexity of the weighting function has been used to formalize ambiguity aversion (see, however, Epstein 1999). Subsequently, convexity was shown to correspond to a pessimistic preference relation under uncertainty (Wakker 2001), where pessimism means that events receive more attention as their associated outcomes are ranked lower. Empirical research has shown, however, that the weighting function does not exhibit pessimism universally. It rather reflects diminishing sensitivity, resulting in an overweighting of unlikely events and an underweighting of likely events (e.g., Tversky and Fox 1995, Wu and Gonzalez 1999, Kilka and Weber 2001). Prominent manifestations of this principle are the possibility effect and the certainty effect, according to which the impact of an event is particularly strong when it turns impossibility into possibility or possibility into certainty. These findings were formalized by means of *bounded subadditivity* (Tversky and Wakker 1995), which comprises *lower subadditivity* (LSA) and *upper subadditivity* (USA). LSA and USA of $W(.)$ are defined as follows:

(3) LSA: $W(A) \geq W(A \cup B) - W(B)$, provided that $W(A \cup B) \leq W(S - E)$ ;

(4) USA: $1 - W(S - A) \geq W(A \cup B) - W(B)$, provided that $W(B) \geq W(E')$ .[1]

Note that LSA of $W(.)$ contradicts convexity.

Under CPT, decision weights are no longer pure measures of belief. The empirical observation that decision makers are less sensitive to uncertainty than to risk motivated Tversky and Fox (1995) to propose a two-stage model in which the decision weight of an event $A$ results from a transformation of its judged probability $q(A)$. When they plotted the

[1] $E$ and $E'$ are boundary events that serve to ensure that the comparisons do not involve both "endpoints" of the family of events (i.e., $\varnothing$ and $S$) simultaneously.

uncertain decision weights against the corresponding judged probabilities, they found that the resulting plots closely resembled the plot of the risky probability weighting function *w*(.) (Fox and See 2003, p. 287). In this two-stage model, the belief component is captured by direct judgment quantifications of degrees of belief, i.e., judged probabilities, that are assumed to satisfy support theory (Tversky and Koehler 1994).

A different approach by Wakker (2004) assumes CPT for prospects of the form $(x, A; 0)$ and $(x, p; 0)$, $x > 0$, and formally shows that the observation that decision makers are less sensitive to uncertainty than to risk is equivalent to the existence of a subadditive and choice-based function $\hat{q}(.)$ such that

(5) $$W(A) = w\,(\,\hat{q}(A)\,)\,,$$

where $\hat{q}(A)$ denotes the choice-based probability of event *A*. It is uniquely determined by $\hat{q}(A) = p$, where *p* is chosen such that indifference between $(x, A; 0)$ and $(x, p; 0)$ holds. From this indifference statement, it follows that the function *w*(.) in Equation (5) is the probability weighting function (for risk). In this decomposition, the belief component can have some source preference in it, i.e., some preference for one source of uncertainty over another.[2]

## 3. Elicitation of Utility Function and Decision Weights

Our approach for eliciting decision weights amounts to a two-step procedure. The first step, which is based on the trade-off method (Wakker and Deneffe 1996), elicits the utility function by determining a standard sequence of outcomes, i.e. a sequence of outcomes equally spaced in utility units. In the second step, decision weights can be computed using the utility values obtained in the first step as inputs.

As in Fennema and van Assen (1999) and Etchart (2003a), we use mixed prospects for the elicitation of the utility function. We obtain standard sequences beginning with the status quo outcome, i.e. $x = 0$. Due to this choice of design, the elicitation of decision weights can be

---

[2] The Ellsberg (1961) paradox represents the most famous illustration of source preference. In this example, one ball is drawn at random from an urn containing 50 red balls and 50 black balls, and from an urn containing 100 red and black balls in an unknown proportion. Ellsberg argued that most people would rather bet on a ball of either color drawn from the first urn than a ball of either color drawn from the second. This corresponds to source preference for the “known” urn. For a formal definition of source preference, see Tversky and Wakker (1995).

The issue of source preference also poses a challenge to the two-stage model of Tversky and Fox (1995); see the discussion in Fox and Tversky (1998), p. 892.

based upon prospects that are particularly easy to compare as they consist of only one non-null outcome.

### *3.1. Elicitation of the Utility Function*

In our experimental investigation, standard sequences are elicited for gains and losses, using monetary outcomes. The elicitation of the standard sequence for gains is performed as follows. Let $R < r < x_0 = 0$ denote three fixed outcomes and $E$ a specified event. As a first step, the outcome $x_1^+$ is determined such that the decision maker is indifferent between the prospects ($x_0$, $E$; $r$) and ($x_1^+$, $E$; $R$). As a second step, the decision maker is called to state the outcome $x_2^+$ such that indifference between the prospects ($x_1^+$, $E$; $r$) and ($x_2^+$, $E$; $R$) holds. Assuming that CPT is an adequate descriptive theory of choice, the combination of the equations resulting from the above two indifference statements implies the equality of $u(x_2^+) - u(x_1^+)$ and $u(x_1^+) - u(x_0)$.

The next steps follow the general principle that once outcome $x_i^+$ has been elicited, outcome $x_{i+1}^+$ leading to indifference between ($x_i^+$, $E$; $r$) and ($x_{i+1}^+$, $E$; $R$) has to be determined. The elicitation procedure results in an increasing sequence of outcomes $x_0$, $x_1^+$, ..., $x_n^+$ such that

(6) $$u(x_{i+1}^+) - u(x_i^+) = u(x_i^+) - u(x_{i-1}^+), \quad i = 1, ..., n-1 .$$

Likewise, with $R' > r' > x_0 = 0$ and the same event $E$, a decreasing standard sequence of losses $x_0$, $x_1^-$, ..., $x_n^-$ is elicited.

### *3.2. Elicitation of Decision Weights*

Let $A_j$ denote an element of the family of events representing the relevant source of uncertainty. The determination of the decision weights for the gain domain, which builds upon the standard sequence of outcomes for gains $x_0 = 0$, $x_1^+$, …, $x_n^+$, proceeds as follows. For each event $A_j$, an outcome $CE_j^+$ is assessed such that the decision maker is indifferent between the prospect ($x_n^+$, $A_j$; 0) and the certain receipt of $CE_j^+$. Under CPT, and with the normalization conditions $u(0) = 0$ and $u(x_n^+) = 1$, this indifference statement translates into:

(7) $$W^+(A_j) = u(CE_j^+) .$$

Using the normalization conditions $u(0) = 0$ and $u(x_n^-) = -1$, decision weights for the loss domain are obtained by means of $W^-(A_j) = -u(CE_j^-)$.[3]

Except by coincidence, $CE_j^\bullet$ itself is not an element of the standard sequence of outcomes, which means that $u(CE_j^\bullet)$ is not immediately available.[4] Because utility is approximately linear over small intervals of outcomes, we determine $u(CE_j^\bullet)$ by linear interpolation between the two adjacent elements of the standard sequence (e.g., Bleichrodt and Pinto 2000). Furthermore, we additionally fit a number of parametric families to our data and use the estimated functions to determine utilities in order to check the robustness of our results. Details will be explained in Subsection 5.2.

### *3.3. Elicitation of Choice-Based Probabilities*

The central idea underlying the two-stage approach is that decision weights comprise a component reflecting belief and a component reflecting decision attitude. In order to make the belief component explicit, a choice-based probability $\hat{q}(A_j)$ is assessed such that the decision maker is indifferent between the risky prospect $(x_n^+, \hat{q}(A_j); 0)$ and the uncertain prospect $(x_n^+, A_j; 0)$, for each event $A_j$. Assuming CPT, this indifference statement implies

(8) $$w^+(\hat{q}(A_j)) = W^+(A_j)\ .$$

It should be noted that the latter equality is derived, without invoking the original two-stage approach, purely by matching a risky simple prospect to an uncertain simple prospect involving the same non-null outcome. The belief component is thus elicited in a "choice-based" manner, as opposed to the approach of asking subjects to state judged probabilities utilized by e.g. Tversky and Fox (1995) or Fox et al. (1996). Equation (8) corresponds to the relation obtained in Theorem 1 of Wakker (2004). Choice-based probabilities can be applied to decompose decision weights for losses according to a similar equation with $w^-(.)$ and $W^-(.)$ instead of $w^+(.)$ and $W^+(.)$.

---

[3] The utility of a loss $x < 0$ is given by $\lambda u(x)$, where $\lambda$ is the loss aversion coefficient. Obviously, this parameter cancels out when calculating decision weights for losses.

[4] Here and henceforth, the superscript $^\bullet$ stands for either $^+$ (gains) or $^-$ (losses).

## 4. Experiment

### *4.1. Subjects and Procedure*

Forty-one subjects (thirty-three male, eight female) participated in our study. All of them were graduate students of business administration at the University of Mannheim. They were enrolled in a decision analysis course and hence familiar with probability and expected utility basics. Each subject received a fixed payment of DM 50 (≈ 25 $) for participation. The experiment took place in February - March 2001.

The experiment was conducted in the form of computer-based individual interview sessions. A special software had been developed for the purpose of the present experiment. The subject and the experimenter were seated in front of a personal computer. The experimenter presented the subject with various choice situations associated with the different experimental tasks and entered the subject's statements into the computer. The presence of the experimenter was meant to help preventing subjects from completing the experimental task without due diligence. Subjects were encouraged to take as much time for reflection as they considered necessary. The mean time needed to complete the experiment as a whole amounted to approximately two hours. Subjects were given the opportunity of a break of a few minutes after the first half of the experiment.

Participants were not directly asked for the specific (outcome or probability) value leading to indifference, i.e. we did not employ a matching procedure. Instead, every value was assessed by means of a series of binary choice questions. The screenshot in Figure 1 in the Appendix displays the typical choice situation that participants faced. If the participant expressed preference for the left-hand (right-hand) prospect, the right-hand prospect was modified to become more (less) attractive. This process continued until the participant regarded the two prospects appearing on the screen as equally attractive. The adjustment process essentially corresponded to successive bisections of the respective (outcome or probability) interval. Once the process had reached the stage that none of the two prospects was strictly preferred, the button "Accept" had to be clicked on. Even after that, a participant noticing a decision error could reenter the adjustment process. Only after clicking on the button "Confirm", the next task appeared on the screen.

### *4.2. Method and Stimuli*

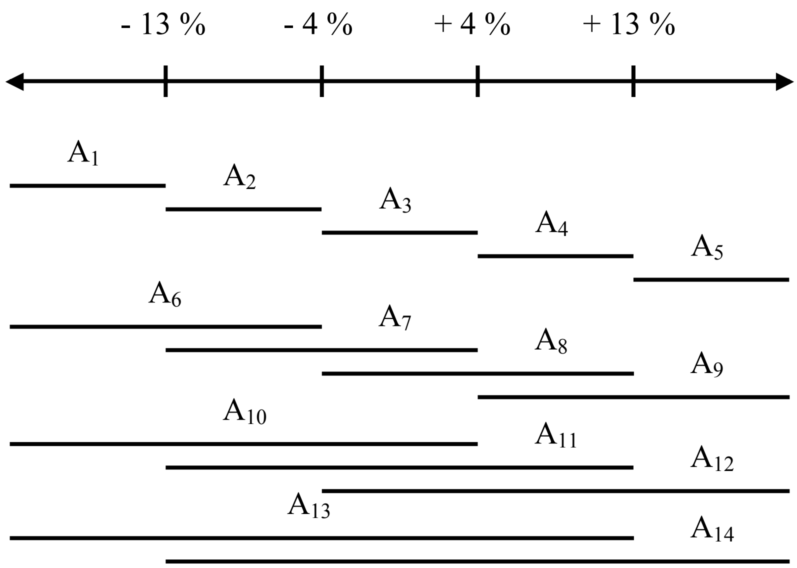
The utility function was elicited from six indifferences. For gains (losses) this amounts to the construction of a standard sequence $x_0 = 0, x_1^+, \ldots, x_6^+$ ($x_0 = 0, x_1^-, \ldots, x_6^-$). We chose $R$

= -400 DM (*R'* = 400 DM) and $r$ = -100 DM (*r'* = 100 DM). The event $E$ was defined as "CDU wins the German general election in 2002". The stimuli were chosen to guarantee that the utility curvature over the interval $[0, x_6^+]$ ($[x_6^-, 0]$) would not be negligible (see Wakker and Deneffe 1996, p. 1139).

The source of uncertainty used to elicit and decompose decision weights was the stock index DAX, which is computed as a capitalization-weighted average of the stock prices of the thirty largest companies listed in Germany. We defined the relevant source of uncertainty as "percentage change of the DAX index over the next six months", measured from the respective day on which the experiment took place.

By the construction of a stock index, the state space is bounded below by -100%, whereas there is no logical upper bound. We partitioned the space of feasible percentage changes to create five elementary events. Abbreviating "percentage change of the DAX index over the next six months" by ΔDAX, these are $A_1$ = {ΔDAX < -13%}, $A_2$ = {-13% ≤ ΔDAX < -4%}, $A_3$ = {-4% ≤ ΔDAX < +4%}, $A_4$ = {+4% ≤ ΔDAX < +13%} and $A_5$ = {+13% ≤ ΔDAX}.[5] Our event space comprises the five elementary events plus all unions formed from the elementary events that result in contiguous intervals. The event space is depicted in Figure 2.

Figure 2: Event space: "percentage change of the DAX index over the next six months"




[5] The rationale behind the chosen partition is as follows. Assuming that the index change roughly follows a normal distribution, we estimated its standard deviation based on current financial data (adjusting for the time horizon in our study) and set the expected change equal to zero (because of the relatively short time horizon in our study). We then fixed interval boundaries such that the cumulative density of each interval amounts to 0.2.

As in Abdellaoui (2000), the elicitation for gains was carried out first. The utility function tasks preceded the decision weights tasks because the latter required $x_6^{\bullet}$ as input. The choice-based probabilities tasks followed the decision weights tasks (for gains). The order in which the events $A_1$, …, $A_{14}$ appeared during the experiment was determined randomly for each subject. It was fixed for all tasks that were based on this set of events (i.e., the decision weights tasks (gains and losses) and the choice-based probabilities task).

In the utility function and the decision weights tasks, outcomes could be varied in steps of 50 DM. Probabilities could be varied in steps of 0.01. The minimum step sizes proved to be sufficiently low so that obtaining two equally preferable prospects was feasible throughout.

In order to be able to assess the reliability of subjects' answers, we presented them twice with several choice situations, or more specifically, with those leading to the determination of $x_1^+$, $x_1^-$, $CE_2^+$, $CE_2^-$, $CE_{12}^+$, $CE_{12}^-$, $\hat{q}_2$, and $\hat{q}_{12}$. These choice situations reappeared at the end of the respective task they belong to except for the outcomes $x_1^+$ and $x_1^-$ which were elicited once again at the end of the whole experiment.

## 5. Results

### *5.1. Reliability*

For purposes of the present study, reliability refers to the stability of subjects' responses when an identical choice task is presented twice. Reliability is measured by means of the Pearson correlation coefficient. Additionally, paired *t* tests are conducted to check for systematic shifts in subjects' responses. Table 1 displays the results.

The overall picture confirms the consistency of subjects' responses. The observed correlation coefficients are quite high, ranging from 0.72 to 0.99. Only for one of the choice-based probabilities tasks ($\hat{q}_{12}$), a significant difference is detected by the paired *t* test ($t_{40}$ = -2.05, $p$ = 0.05; two-tailed).

Table 1: Reliability (two-tailed paired *t* tests and Pearson correlation coefficients)

| | $x_1^+$ | $x_1^-$ | $CE_2^+$ | $CE_2^-$ | $CE_{12}^+$ | $CE_{12}^-$ | $\hat{q}_2$ | $\hat{q}_{12}$ |
|---|---|---|---|---|---|---|---|---|
| $t$ | $t_{29}$ = -1.46 $^{ns}$ | $t_{29}$ = -1.42 $^{ns}$ | $t_{40}$ = 0.40 $^{ns}$ | $t_{40}$ = 0.71 $^{ns}$ | $t_{40}$ = -1.73 $^{ns}$ | $t_{40}$ = 1.60 $^{ns}$ | $t_{40}$ = -0.08 $^{ns}$ | $t_{40}$ = -2.05 $^{*}$ |
| Correlation | 0.90 | 0.75 | 0.92 | 0.95 | 0.79 | 0.99 | 0.72 | 0.78 |

*ns* : non-significant for $\alpha = 0.05$; *: $p < 0.05$

## *5.2. Utility Function*

Both non-parametric and parametric classifications of the shape of the elicited utility functions reveal a predominance of concavity in the gain domain. We obtain some, albeit not very pronounced, evidence in favor of convexity in the loss domain. Our results are quite consistent with previous studies except for a relatively high number of utility functions classified as having a mixed shape through the non-parametric approach. It will be argued below that this finding might be related to the size of the outcome intervals over which the utility functions are constructed.

Table 2: Classification of utility functions (Trade-off method)

| | Gains | | | | Losses | | | |
|---|---|---|---|---|---|---|---|---|
| | concave | linear | convex | mixed | concave | linear | convex | mixed |
| Present study[1] | 46.3% | 7.3% | 19.5% | 26.8% | 22.0% | 22.0% | 24.4% | 31.7% |
| Abdellaoui (2000)[2] | 52.5% | 17.5% | 20.0% | 10.0% | 20.0% | 25.0% | 42.5% | 12.5% |
| Etchart (2003b)[2] | no data | | | | 14.3% | 25.7% | 37.2% | 22.8% |
| Bleichrodt and Pinto (2000)[3] | 59.2% | 26.5% | 2.0% | 12.2% | no data | | | |

[1]: Utility elicited with unknown probabilities; [2]: Utility elicited with known probabilities ;

[3] : Utility elicited with known probabilities and non-monetary outcomes.

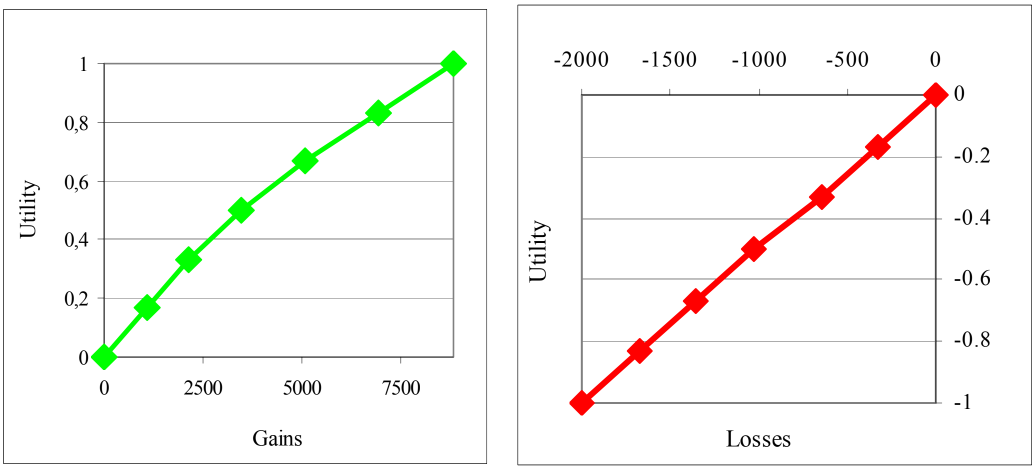
**Non-parametric results.** To classify the curvature of participants' utility functions, we define $d_i^{\bullet} = x_i^{\bullet} - x_{i-1}^{\bullet}$, i = 1, ..., 6, and $\Delta_j^{\bullet} = \text{sign}(d_{j+1}^{\bullet} - d_j^{\bullet})$, j = 1, ..., 5, for both domains separately. A utility function is labeled concave / linear / convex in a particular domain if at least three out of five $\Delta_j^{\bullet}$'s equal 1 / 0 / -1, and mixed otherwise. The idea behind this classification scheme is that even if the utility function of the subject's "true" preference functional (which we assume is CPT) exhibits a unique curvature, the stochastic component of her preference is likely to prevent all $\Delta_j^{\bullet}$'s from being equal. An equivalent classification scheme has been applied in the studies of Abdellaoui (2000), Bleichrodt and Pinto (2000), and Etchart (2003a,b). In contrast to the present study, however, these studies employed risky prospects for the elicitation of the utility function.

The results of the present study are displayed in the first row of Table 2. With respect to the gain domain, 19 subjects (46.3%) are classified as having a concave utility function. Of the other subjects, 3 are classified as linear (7.3%), 8 are classified as convex (19.5%), and 11 are classified as mixed (26.8%). The modal curvature of the utility function in the gain domain is thus concavity, in accordance with both the contention in Tversky and Kahneman (1992) and the respective results of the two related studies. The relative frequency of concave shapes conditional upon a classification as non-mixed amounts to 63.3%, which is quite in line with the corresponding figures of 58.3% in Abdellaoui (2000) and 67.4% in Bleichrodt and Pinto (2000). What distinguishes our findings from the two related studies is the comparatively high percentage of mixed classifications.

With respect to the loss domain, only 10 subjects (24.4%) exhibit the hypothesized convex shape of the utility function. Of the other subjects, 9 are classified as linear (22.0%), 9 are classified as concave (22.0%), and 13 are classified as mixed (31.7%). The preponderance of mixed shapes parallels the earlier observation of the relatively high number of mixed classifications in the gain domain. The lack of a distinct pattern among the non-mixed classifications might be due to the fact that the average size of the outcome interval $[x_6^-, 0]$ is relatively small. This conjecture is consistent with a considerable body of evidence suggesting that the utility function is approximately linear over small and moderate stakes (see, e.g., Lopes and Oden 1999, Rabin 2000). Interestingly, Etchart (2003b) found that the percentage of mixed cases declines in favor of concavity when the average size of the loss interval under investigation is higher (22.5% mixed cases for small losses as given in Table 2, and 11.5%

mixed cases for higher losses)[6]. The percentages of linear and convex cases remained constant. Moreover, both in Abdellaoui (2000) and in Fennema and van Assen (1999) the evidence in favor of convexity in the loss domain is somewhat weaker than the evidence in favor of concavity in the gain domain.

Figure 3: Utility functions (for mean values)




When only concave and convex shapes are taken into account, a binomial test reveals that there are significantly more concave utility functions in the gain domain than convex ones ($p = 0.03$, one-tailed), whereas the reverse statement for the loss domain cannot be established ($p = 0.5$, one-tailed). These findings are also mirrored in the aggregate data. Figure 3 displays the utility functions, separately for each domain, that result when the elements of the respective standard sequences are averaged over participants.

**Parametric results.** In addition to the non-parametric classification of utility functions presented above, parametric estimations were conducted using nonlinear least squares fits. Their purpose is twofold: First, they provide another way of looking at the utility function data and allow for comparisons with studies that adopt parametric approaches.

---

[6] In Etchart (2003b), the mean interval for small losses is [FF -8,460; FF -1,500]; the mean interval for high losses is [FF -100,000; FF -20,000] (1 US$ ≈ 6 FF).

Second, the estimated utility functions serve as input in the determination of decision weights as explained in Section 3. Two parametric forms were used: power and exponential (Table 3).

Table 3: Parametric specifications

| | Power | Exponential |
|---|---|---|
| Gains | $z^{\alpha+}$ | $(1-exp(-\lambda^{+}z))/ (1-exp(-\lambda^{+}))$ |
| Losses | $-(-z)^{\alpha-}$ | $-(1-exp(\lambda^{-}z))/(1-exp(-\lambda^{-}))$ |

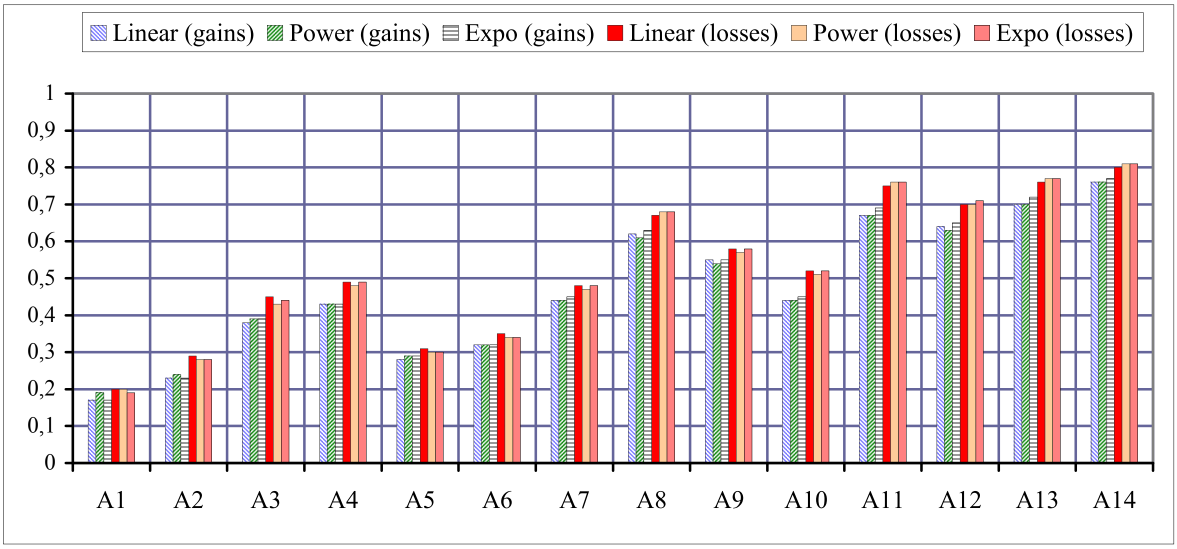
The power family is frequently employed in experimental studies on utility measurement. Its widespread use possibly has roots in the power law of psychophysics (Stevens 1957); for an early application in risky decision making see Tversky (1967). Exhibiting the simple property of constant absolute risk aversion, the exponential specification is also extensively used in utility assessment from laboratory experiments as well as decision analysis interviews. The domain $[0, x_6^{+}]$ for gains ($[x_6^{-}, 0]$ for losses) is mapped into the positive (negative) unit interval through the rescaling $z = x/x_6^{+}$ ($z = -x/x_6^{-}$).

Table 4: Summary statistics for parameters of the utility function

| | Gains | | | Losses | | |
|---|---|---|---|---|---|---|
| | median | mean | st.dev. | median | mean | st.dev. |
| *Power:* $\alpha^{\bullet}$ | 0.91 | 0.91 | 0.32 | 0.96 | 1.08 | 0.47 |
| *Exponential:* $\lambda^{\bullet}$ | 0.28 | 0.57 | 1.21 | 0.09 | 0.03 | 1.19 |

For gains, both the power utility function and the exponential utility function exhibit a concave shape for the mean and median parameter estimates, so that the non-parametric findings are confirmed. At an individual level, both power and exponential estimates show that the predominant shape is concavity (26 out of 41 subjects satisfy $\alpha^{+} < 1$ and $\lambda^{+} > 0$ respectively). The estimate for the power utility function almost coincides with the median estimate of 0.89 in Abdellaoui (2000) and the median estimate of 0.88 in Tversky and Kahneman (1992). With regard to the loss domain, the parametric fitting also supports the non-parametric results that were at variance with the hypothesized convex shape. For the power utility function and the exponential utility function, the mean and median parameter estimates are near to or even on either side of the respective borderline separating concave and convex shapes. At an individual level, convexity for losses seems less dominating than concavity for gains (23 out of 41 utility functions are classified as convex). The parameter estimate for the power utility function therefore deviates from the median estimate of 0.92 reported in Abdellaoui (2000) and the median estimate of 0.88 obtained by Tversky and Kahneman (1992).

Figure 4: Mean decision weights




*5.3. Decision Weights*

Decision weights conform to the property of bounded subadditivity almost without exception, which strongly questions the interpretation of subjects' behavior as "SEU plus noise". Whereas the degree of subadditivity does not differ systematically between the gain and loss domains, decision weights appear to exhibit higher elevation for losses. We do not find evidence for a systematic violation of the duality condition. The aforementioned results do not critically depend upon the particular interpolation method used for the utility function.

**Decision weights for gains and losses.** As explained in Subsection 3.2, the decision weight of event $A_j$, $W^+(A_j)$ ($W^-(A_j)$), equals the (absolute value of the) utility of the certainty equivalent of the prospect $(x_n^+, A_j; 0)$ ($(x_n^-, A_j; 0)$). The utility function is obtained by both linear interpolation between the outcomes of the standard sequence and (individual) parametric fitting of the families listed in Table 3, separately for each subject.

Figure 4 displays mean decision weights for gains and losses. It can be seen that mean decision weights satisfy all monotonicity conditions imposed by the structure of the event space as depicted in Figure 2. Overall, Figure 4 suggests that, for each outcome domain, mean decision weights are not sensitive to the parametric specification used for the utility function. One-factor ANOVA tests with repeated measures detect significant differences between the three specifications in 11 (5) out of 28 cases for α fixed at 0.05 (0.01). It seems, however, that the "traditional" power family produces decision weights that are closer to those computed by means of linear interpolation. This is confirmed by (two-tailed) paired *t* tests which detect significant differences between linear interpolation and the power specification in only 4 (2) out of 28 cases if α is fixed at 0.05 (0.01).

Table 5: Elevation of decision weights: $W^{+}(A)$ vs. $W^{-}(A)$ (paired *t* tests, one-tailed)

| | Linear interpolation | Power approximation | Expo. approximation |
|---|---|---|---|
| | $t_{40}$ | $t_{40}$ | $t_{40}$ |
| $A_1$ | 0.72 *ns* | 0.26 *ns* | 0.48 *ns* |
| $A_2$ | 2.26 * | 1.75 * | 1.80 * |
| $A_3$ | 2.13 * | 1.39 *ns* | 1.51 *ns* |
| $A_4$ | 1.67 *ns* | 1.36 *ns* | 1.39 *ns* |
| $A_5$ | 0.65 *ns* | 0.05 *ns* | 0.15 *ns* |
| $A_6$ | 1.02 *ns* | 0.64 *ns* | 0.52 *ns* |
| $A_7$ | 1.25 *ns* | 0.96 *ns* | 0.87 *ns* |
| $A_8$ | 1.81 * | 2.25 * | 1.83 * |
| $A_9$ | 1.00 *ns* | 1.04 *ns* | 0.87 *ns* |
| $A_{10}$ | 2.23 * | 2.20 * | 1.99 * |
| $A_{11}$ | 2.13 * | 2.42 * | 2.00 * |
| $A_{12}$ | 1.62 *ns* | 2.17 * | 1.83 * |
| $A_{13}$ | 1.83 * | 2.21 * | 1.74 * |
| $A_{14}$ | 1.29 *ns* | 1.70 * | 1.14 *ns* |

*ns*: non-significant for α = 0.05; *: $p < 0.05$

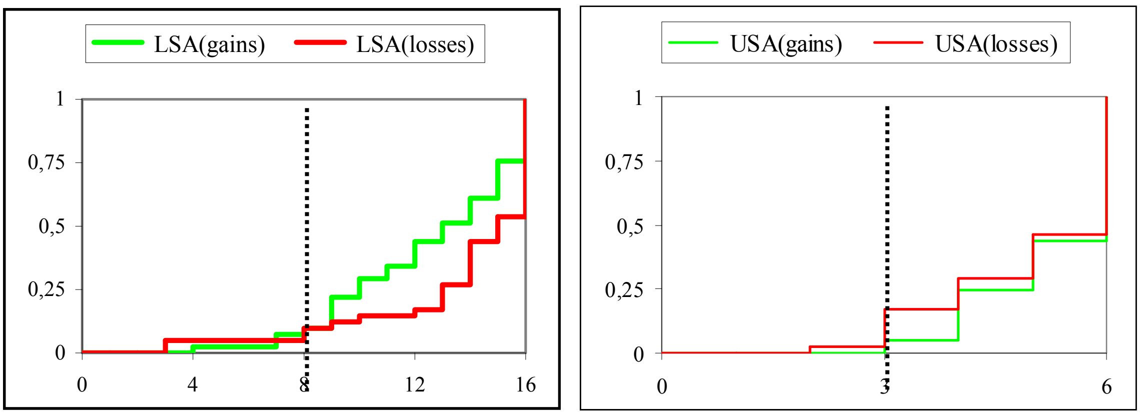
Figure 4 also shows that mean decision weights referring to the loss domain exceed their gain domain counterparts for all events. In the terminology used for the probability weighting function under risk (Gonzalez and Wu 1999), decision weights for the loss domain seem to exhibit more elevation. This statement is statistically supported more for likely events than for unlikely ones. As can be seen from Table 5, paired *t* tests, conducted separately for each event, lead to a rejection of the null hypothesis $W^+(A_j) = W^-(A_j)$ in favor of $W^+(A_j) < W^-(A_j)$ in about one half of the 14 cases ($\alpha = 0.05$, one-tailed).

It should be observed, however, that the event-wise comparison of decision weights is not necessarily a very effective technique to detect potential differences between the two domains. This limitation can be overcome using the two-stage approach that permits a characterization of the elevation of a participant's decision weights by means of a single overall parameter. We will therefore return to the issue of elevation in Subsection 5.5.

**Sensitivity of elicited decisions weights.** The notion of sensitivity is related to the bounded subadditivity property of decision weights presented in Section 2 which says that the impact of an event is particularly strong if it is added to the null event or subtracted from the universal event. The structure of the event space in the present study leads to 16 conditions to test for LSA[7] and to 6 conditions to test for USA[8], separately for each domain. For each participant, the respective number of conditions satisfied is computed. To obtain conservative results, a condition counts as satisfied only if strict inequality holds. Figure 5 displays the empirical distribution functions of conditions satisfied.

Figure 5: Empirical distributions of subadditivity conditions satisfied

(Linear interpolation)




[7] LSA implies e.g. that $W^\bullet(A_1) \geq W^\bullet(A_6) - W^\bullet(A_2)$, where $A_6 = A_1 \cup A_2$. The fifteen other triples of events ($A_i$; $A_i \cup A_j$, $A_j$) allowing a test of LSA are ($A_2$; $A_7$, $A_3$), ($A_3$; $A_8$, $A_4$), ($A_4$; $A_9$, $A_5$), ($A_1$; $A_{10}$, $A_7$), ($A_2$; $A_{11}$, $A_8$), ($A_3$; $A_{10}$, $A_6$), ($A_3$; $A_{12}$, $A_9$), ($A_4$; $A_{11}$, $A_7$), ($A_5$; $A_{12}$, $A_8$), ($A_1$; $A_{13}$, $A_{11}$), ($A_2$; $A_{14}$, $A_{12}$), ($A_4$; $A_{13}$, $A_{10}$), ($A_5$; $A_{14}$, $A_{11}$), ($A_6$; $A_{13}$, $A_8$), and ($A_7$; $A_{14}$, $A_9$). All the triples satisfy the boundary condition $W^\bullet(A_i \cup A_j) \leq W^\bullet(S - E)$ for some non-null event $E$.

[8] USA implies e.g. that $1 - W^\bullet(A_{14}) \geq W^\bullet(A_6) - W^\bullet(A_2)$, where $S - A_{14} = A_6 - A_2 = A_1$. The five other triples of events ($S - A_i$; $A_i \cup A_j$, $A_j$) allowing a test of USA are ($A_{14}$; $A_{10}$, $A_7$), ($A_{14}$; $A_{13}$, $A_{11}$), ($A_{13}$; $A_9$, $A_4$), ($A_{13}$; $A_{12}$, $A_8$), and ($A_{12}$; $A_{10}$, $A_3$). All the triples satisfy the boundary condition $W^\bullet(A_j) \geq W^\bullet(E')$ for some non-null event $E'$.

If decision weights were perfectly additive, the number of conditions satisfied would equal zero in all cases. If decision weights were fundamentally additive but with a nonsystematic error component, the number of conditions satisfied would be stochastic and its distribution would be centered around one half of the number of available conditions.

Figure 5 provides strong evidence against "noisy additivity" of decision weights and, therefore, against the descriptive validity of SEU. With respect to LSA (USA), more than 8 (3) conditions are satisfied by 90.2% (95.1%) of the subjects in the gain domain and by 90.2% (82.9%) of the subjects in the loss domain. Aggregating over all subjects, the proportion of LSA (USA) conditions satisfied amounts to 78.7% (87.8%) in the gain domain and to 86.4% (84.1%) in the loss domain. These figures closely resemble the results reported in Tversky and Fox (1995), Table 4 (p. 276). The median number of LSA (USA) conditions satisfied is 13 (6) in the gain domain and 15 (6) in the loss domain.

Table 6: Tests of duality (*t* tests, two-tailed)

| | Linear interpolation | | Power approximation | | Exponential approximation | |
|---|---|---|---|---|---|---|
| | mean | $t_{40}$ | mean | $t_{40}$ | mean | $t_{40}$ |
| $W^-(A_1) + W^+(S - A_1)$ | 0.955 | -1.42 $^{ns}$ | 0.954 | -1.41 $^{ns}$ | 0.965 | -1.09 $^{ns}$ |
| $W^-(A_5) + W^+(S - A_5)$ | 1.012 | 0.29 $^{ns}$ | 0.997 | -0.08 $^{ns}$ | 1.012 | 0.30 $^{ns}$ |
| $W^-(A_6) + W^+(S - A_6)$ | 0.989 | -0.28 $^{ns}$ | 0.971 | -0.74 $^{ns}$ | 0.985 | -0.39 $^{ns}$ |
| $W^-(A_9) + W^+(S - A_9)$ | 1.020 | 0.48 $^{ns}$ | 1.013 | 0.33 $^{ns}$ | 1.032 | 0.74 $^{ns}$ |
| $W^-(A_{10}) + W^+(S - A_{10})$ | 1.067 | 1.62 $^{ns}$ | 1.054 | 1.25 $^{ns}$ | 1.073 | 1.63 $^{ns}$ |
| $W^-(A_{12}) + W^+(S - A_{12})$ | 1.012 | 0.32 $^{ns}$ | 1.023 | 0.69 $^{ns}$ | 1.032 | 0.88 $^{ns}$ |
| $W^-(A_{13}) + W^+(S - A_{13})$ | 1.044 | 1.30 $^{ns}$ | 1.065 | 1.90 $^{ns}$ | 1.064 | 1.77 $^{ns}$ |
| $W^-(A_{14}) + W^+(S - A_{14})$ | 0.973 | -0.80 $^{ns}$ | 0.998 | -0.05 $^{ns}$ | 0.985 | -0.42 $^{ns}$ |

*ns*: non-significant for $\alpha = 0.05$

To also obtain a quantification of the degree of subadditivity of decision weights, the measure initially applied by Tversky and Fox (1995) and thereafter used by for instance Kilka and Weber (2001) is computed. An equivalent statement of the condition defining LSA (USA) is $W^\bullet(A_i) + W^\bullet(A_j) - W^\bullet(A_i \cup A_j) \geq 0$ $(1 - W^\bullet(S - A_i) - (W^\bullet(A_i \cup A_j) - W^\bullet(A_j)) \geq 0)$. The index of LSA (USA), henceforth denoted by $\mu_W^{\bullet}$(LSA) ($\mu_W^{\bullet}$(USA)), is computed as the mean value of the left-hand side of the above inequalities, taken over the 16 (6) triples of events allowing a test of the respective condition. It is computed separately for each participant and each domain. The median values (where medians are taken over participants) of the indices amount to $\mu_W^{+}$(LSA) = 0.18 ($\mu_W^{+}$(USA) = 0.20) in the gain domain and to $\mu_W^{-}$(LSA) = 0.25 ($\mu_W^{-}$(USA) = 0.18) in the loss domain. The degree of subadditivity in our data is consistent with the findings of Tversky and Fox (1995), Table 5 (p. 277), and with the findings concerning the source of uncertainty with higher perceived competence in Kilka and Weber (2001), Table 1 (p. 1719). The pervasiveness of subadditivity is also mirrored by the fact that the LSA (USA) index is strictly positive for 90.2% (100%) of the subjects in the gain domain and for 92.7% (92.7%) of the subjects in the loss domain. Both *t* tests of the hypothesis that the mean index value is zero and sign tests of the hypothesis that positive and negative index values are equally likely result in $p < 0.001$ for each of the four domain-property-combinations.

We also considered whether the degree of subadditivity of decision weights differs across the two domains. The null hypothesis of equal mean (respectively median) LSA / USA index values cannot be rejected by either paired *t* tests or sign tests at conventional levels of significance (each $p > 0.05$, one-tailed). The null hypothesis that the mean LSA (USA) index value is equal for gains and losses cannot be rejected by either paired *t* tests or sign tests at conventional levels of significance (each $p > 0.05$, one-tailed). This result suggests that the sensitivity of decision weights does not exhibit domain-dependence. An alternative characterization of decision weights in terms of elevation and sensitivity will be presented in Subsection 5.5.

**Duality of elicited decision weights.** If subjects' decision weights systematically violate duality, one of the additional degrees of freedom introduced by CPT (as compared to CEU) is empirically warranted. Owing to the structure of the event space in the present study, there are eight pairs of events and complementary events that provide a test of the duality condition. Table 6 displays mean values of the respective sums of decision weights. Additionally, it shows *t* values for tests of the hypothesis that mean sums equal 1.

The overall picture is remarkably clear. The respective duality conditions are not significantly violated ($\alpha = 0.05$, two-tailed).

**Error propagation.** Because the trade-off method is "chained", i.e. the elicitation of $x_i$ requires $x_{i-1}$ as input, errors may propagate and produce "noisy" decision weights. To check the impact of error propagation, we performed a simulation based on an error theory comparable to Hey and Orme (1994). More specifically, we assumed that the response error ε is a proportion of the assessed utility difference. In other terms, it is equal to (1+ε) times the true utility difference as in Bleichrodt and Pinto (2000). The error term was assumed to be normally distributed, with mean 0 and standard deviation 0.05.

Overall, about 1,600 simulations using individual subject data were performed, resulting in simulated decision weights for gains and losses (obtained through linear interpolation). For each event, the proportion of the mean absolute differences between simulated and observed decision weights to the mean observed decision weights do not exceed 3.5% for gains as well as for losses. This proportion does not exceed 4.5% when mean absolute differences are replaced by standard deviations. To sum up, as in Bleichrodt and Pinto (2000), the simulation shows a rather low impact of error propagation in the tradeoff method on decision weights.

Table 7: Summary statistics for choice-based probabilities

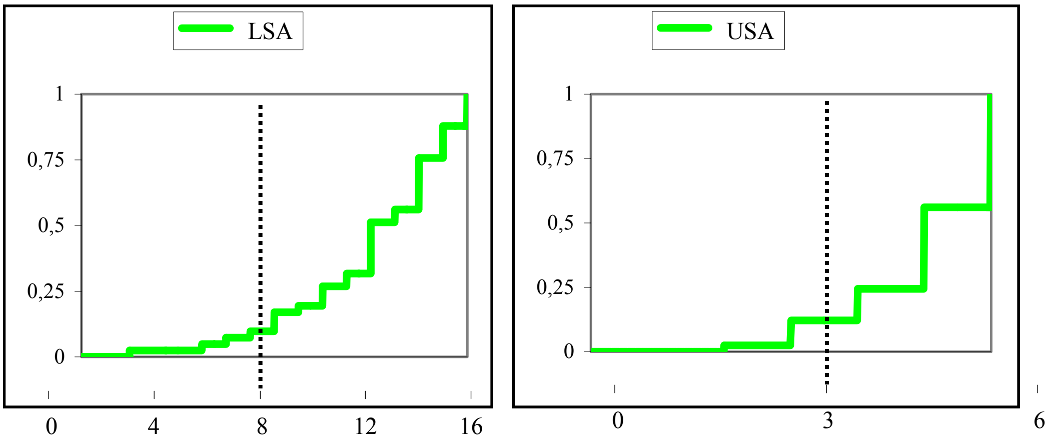
| | $\hat{q}_1$ | $\hat{q}_2$ | $\hat{q}_3$ | $\hat{q}_4$ | $\hat{q}_5$ | $\hat{q}_6$ | $\hat{q}_7$ | $\hat{q}_8$ | $\hat{q}_9$ | $\hat{q}_{10}$ | $\hat{q}_{11}$ | $\hat{q}_{12}$ | $\hat{q}_{13}$ | $\hat{q}_{14}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| median | 0.10 | 0.17 | 0.35 | 0.42 | 0.20 | 0.25 | 0.40 | 0.66 | 0.55 | 0.50 | 0.75 | 0.72 | 0.75 | 0.85 |
| mean | 0.13 | 0.20 | 0.36 | 0.43 | 0.23 | 0.27 | 0.42 | 0.62 | 0.54 | 0.48 | 0.73 | 0.68 | 0.75 | 0.82 |
| st.dev. | 0.09 | 0.13 | 0.19 | 0.21 | 0.15 | 0.15 | 0.19 | 0.20 | 0.21 | 0.17 | 0.16 | 0.17 | 0.13 | 0.13 |

### *5.4. Probabilities and Decision Weights*

As was found in previous experimental investigations for judged probabilities, the elicited degrees of belief exhibit bounded subadditivity. The degree of subadditivity is lower than for decision weights, which is also in line with previous research. The comparison of decision weights and choice-based probabilities reveals an overweighting of unlikely events for both gains and losses. An underweighting of likely events occurs for gains but not for losses.

**Choice-based probabilities.** Table 7 contains summary statistics (median, mean, and standard deviation) for the choice-based probabilities. Parallel to the finding for decision weights, it can be verified that median and mean probabilities satisfy all monotonicity conditions imposed by the structure of the event space.[9]

Figure 6: Subadditivity of choice-based probabilities




**Subadditivity of choice-based probabilities.** As pointed out in detail in Section 2 and Subsection 3.3, degrees of belief in the present study are determined through choices and in this respect differ from judged probabilities elicited in studies relying on the original two-stage model. This aspect naturally leads to the question of whether the properties of the elicited degrees of belief are affected by the elicitation mode. Subadditivity is one of the key properties of judged probabilities in support theory (Tversky and Koehler 1994) and also crucial in the derivation of the decomposition in Wakker (2004). Parallel to the analysis of the subadditivity property for decision weights, Figure 6 displays the empirical distribution functions of LSA / USA conditions satisfied.

The graphs in Figure 6 closely resemble their decision weight counterparts in Figure 5. With respect to LSA (USA) of choice-based probabilities, more than 8 (3) conditions are satisfied by 82.9% (87.8%) of the subjects. We also computed indices of subadditivity for choice-based probabilities, denoted by $\mu_q$(LSA) and $\mu_q$(USA), which are defined analogously

---

[9] Assuming the two-stage decomposition $W^{\bullet}(A_i) = w^{\bullet}(\hat{q}(A_i))$, the monotonicity condition for subjective choice-based probabilities, i.e. $\hat{q}(A) \leq \hat{q}(B)$ for all events $A$, $B$ with $A \subset B$, follows from the corresponding condition for decision weights and the fact that the transformation function $w^{\bullet}(.)$ is increasing.

to the respective measures for decision weights, $\mu_W$(LSA) and $\mu_W$(USA), presented in Subsection 5.3. The LSA (USA) index is strictly positive for 90.2% (97.6% ) of the subjects. The median values (where medians are taken over participants) of these indices amount to $\mu_q$(LSA) = 0.12 and $\mu_q$(USA) = 0.13. Two conclusions can be drawn from these figures. First, they provide strong empirical support for subadditivity of choice-based probabilities, which thereby share an important property with judged probabilities (Tversky and Koehler 1994, Tversky and Fox 1995, Fox et al. 1996, Wu and Gonzalez 1999, Kilka and Weber 2001). Second, since the indices of subadditivity are smaller for probabilities than for decision weights, our data are consistent with an interpretation of decision weights as a subadditive measure of belief transformed by a subadditive function (Fox and Tversky 1998, Wakker 2003).

Table 8: Decision weights vs. choice-based probabilities (paired *t* tests, one-tailed)

| | Linear interpolation | | Power approximation | | Expo. approximation | |
|---|---|---|---|---|---|---|
| | Gains | Losses | Gains | Losses | Gains | Losses |
| | $t_{40}$ | $t_{40}$ | $t_{40}$ | $t_{40}$ | $t_{40}$ | $t_{40}$ |
| $W^{\bullet}(A_1) - \hat{q}_1$ | 1.51 $^{ns}$ | 3.92 $^{***}$ | 2.06 $^{*}$ | 3.63 $^{***}$ | 1.56 $^{ns}$ | 3.63 $^{***}$ |
| $W^{\bullet}(A_2) - \hat{q}_2$ | 1.56 $^{ns}$ | 3.63 $^{***}$ | 1.40 $^{ns}$ | 4.17 $^{***}$ | 1.06 $^{ns}$ | 4.09 $^{***}$ |
| $W^{\bullet}(A_3) - \hat{q}_3$ | 1.06 $^{ns.}$ | 4.09 $^{***}$ | 0.87 $^{ns}$ | 3.06 $^{**}$ | 0.92 $^{ns}$ | 3.46 $^{***}$ |
| $W^{\bullet}(A_4) - \hat{q}_4$ | 0.92 $^{ns}$ | 3.46 $^{***}$ | -0.06 $^{ns}$ | 1.64 $^{n.s.}$ | 0.05 $^{ns}$ | 1.79 $^{*}$ |
| $W^{\bullet}(A_5) - \hat{q}_5$ | 0.05 $^{ns}$ | 1.79 $^{*}$ | 2.20 $^{*}$ | 3.18 $^{**}$ | 1.99 $^{*}$ | 3.20 $^{**}$ |
| $W^{\bullet}(A_6) - \hat{q}_6$ | 1.99 $^{*}$ | 3.20 $^{**}$ | 2.06 $^{*}$ | 2.89 $^{**}$ | 2.05 $^{*}$ | 2.98 $^{**}$ |
| $W^{\bullet}(A_7) - \hat{q}_7$ | 2.05 $^{*}$ | 2.98 $^{**}$ | 0.76 $^{ns}$ | 2.08 $^{*}$ | 1.00 $^{ns}$ | 2.29 $^{*}$ |
| $W^{\bullet}(A_8) - \hat{q}_8$ | 1.00 $^{ns}$ | 2.29 $^{*}$ | -0.20 $^{ns}$ | 2.87 $^{**}$ | 0.27 $^{ns}$ | 2.89 $^{**}$ |
| $W^{\bullet}(A_9) - \hat{q}_9$ | 0.27 $^{ns}$ | 2.89 $^{**}$ | -0.20 $^{ns}$ | 0.97 $^{ns}$ | 0.25 $^{ns}$ | 1.20 $^{ns}$ |
| $W^{\bullet}(A_{10}) - \hat{q}_{10}$ | 0.25 $^{ns}$ | 1.20 $^{ns}$ | -1.54 $^{ns}$ | 0.97 $^{ns}$ | -1.14 $^{ns}$ | 1.21 $^{ns}$ |
| $W^{\bullet}(A_{11}) - \hat{q}_{11}$ | -1.14 $^{ns}$ | 1.21 $^{ns}$ | -1.81 $^{*}$ | 1.11 $^{ns}$ | -1.33 $^{ns}$ | 1.17 $^{ns}$ |
| $W^{\bullet}(A_{12}) - \hat{q}_{12}$ | -1.33 $^{ns}$ | 1.17 $^{ns}$ | -1.44 $^{ns}$ | 1.31 $^{ns}$ | -0.98 $^{ns}$ | 1.43 $^{ns.}$ |
| $W^{\bullet}(A_{13}) - \hat{q}_{13}$ | -0.98 $^{ns}$ | 1.43 $^{ns}$ | -1.82 $^{*}$ | 1.28 $^{ns}$ | -1.18 $^{ns}$ | 1.41 $^{ns.}$ |
| $W^{\bullet}(A_{14}) - \hat{q}_{14}$ | -1.18 $^{ns}$ | 1.41 $^{ns}$ | -2.12 $^{*}$ | -0.71 $^{ns}$ | -1.57 $^{ns}$ | -0.68 $^{ns}$ |

*ns*: non-significant for $\alpha = 0.05$; *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$

**Decision weights versus choice-based probabilities.** It was already mentioned that CEU coincides with SEU for the special case that decision weights are additive subjective probabilities. Contrasting decision weights and choice-based probabilities, it can be asked whether they coincide for every event considered. Table 8 displays the results of paired *t* tests conducted to check whether mean decision weights equal mean choice-based probabilities, separately for each event and each domain.

The pattern in the gain domain reflects an overweighting of unlikely events and an underweighting of likely events (relative to the respective choice-based probabilities). This finding is broadly consistent with evidence from studies of decision making under risk, particularly an inverse S-shaped transformation function (see e.g. Tversky and Kahneman 1992, Wu and Gonzalez 1996, Bleichrodt and Pinto 2000). The statistical significance of the observed differences is not high, though. The pattern in the loss domain, which is foreshadowed by the pattern in the gain domain and the comparison of decision weights across domains presented in Figure 4, indicates a marked overweighting of unlikely events. An underweighting of likely events cannot be found in the data, yet the decrease of the difference between decision weights and choice-based probabilities as one moves towards more likely events is common to both domains. From an overall perspective, the results contained in Table 8 suggest that CPT is descriptively more adequate than SEU.

### *5.5. Fitting the two-stage model*

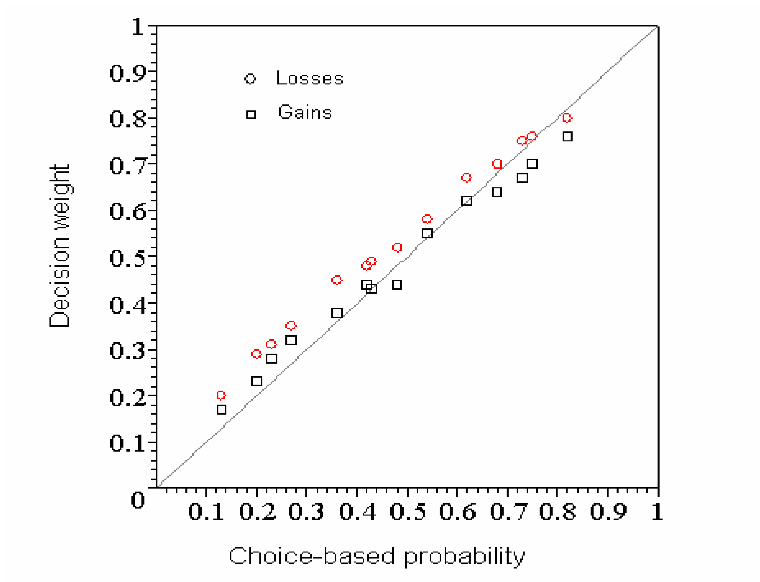
Fitting the two-stage model to our data amounts to estimating the parameter(s) of the probability weighting function via nonlinear regression with decision weights as dependent variable and choice-based probabilities as explanatory variable (Figure 7 plots mean choice-based probabilities against mean decision weights). We focus on parametric specifications of the probability weighting function that permit a clear separation between elevation and curvature. Whereas the curvature parameter does not differ significantly between the gain and loss domains, the elevation parameter is markedly higher for losses. The results obtained under the two-stage model therefore mirror their counterparts from the holistic analysis of decision weights in Subsection 5.3.

The estimations are conducted separately for each participant. We employ several parametric forms originally proposed for the probability weighting function under risk. One parametric specification that proves to be particularly useful is the linear-in-log-odds form applied by for instance Goldstein and Einhorn (1987), Lattimore et al. (1992) and Gonzalez and Wu (1999):

(9) $$w(p) = \frac{\delta \cdot p^{\gamma}}{\delta \cdot p^{\gamma} + (1-p)^{\gamma}}.$$[10]

Its usefulness derives from the fact that it permits a distinction between two essential features of the probability weighting function: elevation and curvature (i.e., sensitivity in the context of probability weighting). For the linear-in-log-odds specification, the parameter δ mainly controls elevation, whereas the parameter γ mainly controls curvature. The issues of elevation and curvature of decision weights addressed in Subsection 5.3 can therefore be reanalyzed in the light of the respective parameter estimates of the probability weighting function in the two-stage model.[11]

Figure 7: Mean decision weights as a function of choice-based probability




[10] The linear-in-log-odds property is demonstrated in Gonzalez and Wu (1999), p. 139.

[11] The same kind of decomposition is feasible for the (two-parameter) probability weighting function derived axiomatically by Prelec (1998): $w(p) = exp(-\beta \cdot (-log(p))^{\alpha})$, where the parameter β mainly controls elevation and the parameter α mainly controls curvature. We also applied this probability weighting function in the analyses to be presented subsequently. For a small number of subjects, we obtained parameter estimates – especially for the parameter β – that were quite high, which would have unduly affected summary statistics and statistical inference. We therefore focus on the linear-in-log-odds function, complementing it with a parsimonious and particularly robust specification of the probability weighting function to avoid dependence upon a single parametric form.

A parsimonious parametric form for the probability weighting function that nevertheless incorporates a clear separation between elevation and curvature is the linear approximation

(10) $w(p) = \alpha + \beta \cdot p \quad for \; p \in (0,1), w(0) = 0, w(1) = 1$ ,

applied for instance in Wu and Gonzalez (1996) and Kilka and Weber (2001). Curvature is captured by the parameter β, whereas elevation can be characterized through $\int w(p)\mathrm{d}p$.[12]

By construction, probability weighting functions with a single free parameter like the one used by Karmarkar (1978), Tversky and Kahneman (1992) or the one-parameter variant in Prelec (1998) do not permit an independent variation of elevation and curvature. For this reason, the following presentation of results is restricted to the two-parameter specifications.

Table 9 conveys parameter estimates of the probability weighting function for median data resulting from nonlinear least squares with a normally distributed error term.[13] It can be seen that, for linear interpolation as well as power or exponential approximations, the curvature parameter for median data is quite similar across the two domains for all parametric specifications of the probability weighting function. In contrast, the estimates of δ show that the probability weighting function for median data exhibit more elevation for losses than for gains.[14] These estimates do not seem however very close to those obtained for decision making under risk in Abdellaoui (2000) for gains and losses ($\delta = 0.65$ and $\gamma = 0.60$ for gains; $\delta = 0.84$ and $\gamma = 0.65$ for losses) and in Gonzalez and Wu (1999) for gains ($\delta = 0.77$ and $\gamma = 0.44$)[15]. In decision making under risk, higher elevation of the probability weighting function for losses has been found by both Abdellaoui (2000) and Wu et al. (2003) in their re-analysis of the Tversky and Kahneman (1992) data using a two-parameter function.

---

[12] Bounded subadditivity of the probability weighting function corresponds to $\alpha > 0$ and $\alpha + \beta < 1$, in which case the measure of elevation equals $\alpha + \beta/2$ (see also Kilka and Weber 2001, p. 1717).

[13] We also conducted the nonlinear regressions using individual subject data. The median (across subjects) estimated parameter values are quite in line with the results presented in Table 9.

[14] The quality of fit is systematically higher under the linear-in-log-odds specification, for all domains and all kinds of utility interpolation. The mean adjusted $R^2$ ranges from 0.94 to 0.97 for the linear-in-log-odds specification and from 0.74 to 0.83 for the linear approximation.

Table 9: Parameter estimates of the probability weighting function (median data)

| | | Linear-in-log-odds | | Linear weighting | |
|---|---|---|---|---|---|
| | | elevation : δ | curvature : γ | elevation | Curvature : β |
| Linear interpolation | Gains | 0.987 (0.04) | 0.860 ( 0.04) | 0.497 (0.008) | 0.889 (0.03) |
| | Losses | 1.277 ( 0.06) | 0.786 ( 0.05) | 0.551 (0.011) | 0.838 (0.04) |
| Power approximation | Gains | 0.975 ( 0.04) | 0.832 (0.04) | 0.495 (0.009) | 0.865 (0.03) |
| | Losses | 1.345 (0.05) | 0.842 (0.03) | 0.561 (0.007) | 0.884 (0.03) |
| Exponential approximation | Gains | 0.981 ( 0.04) | 0.907 (0.05) | 0.498 (0.009) | 0.922 (0.038) |
| | Losses | 1.318 (0.05) | 0.866 (0.04) | 0.556 (0.008) | 0.905 (0.03) |

*Note.* Values in parentheses are standard errors.

At the level of individual subjects, paired *t* tests are conducted to investigate the hypothesis of equal parameter values in the gain and loss domains. Whereas the curvature parameter (γ) is statistically indistinguishable across domains at conventional levels of significance ($t_{40} = 0.00$, $p = 1.00$ for the linear interpolation; $t_{40} = 1.26$, $p = 0.21$ for the power approximation; $t_{40} = 1.18$, $p = 0.25$ for the exponential approximation; two-tailed), the domain-dependence of the elevation parameter (δ) is confirmed ($t_{40} = 2.04$, $p = 0.02$ for the linear interpolation; $t_{40} = 2.08$, $p = 0.02$ for the power approximation; $t_{40} = 1.23$, $p = 0.11$ for the exponential approximation; one-tailed).

The conclusions derived for the linear-in-log-odds specification turn out to be robust when the simple linear probability weighting function is considered instead. Paired *t* tests show no significant difference between the curvature parameter (β) of the gain and loss domains ($t_{40} = 1.12$, $p = 0.27$ for the linear interpolation; $t_{40} = 1.55$, $p = 0.13$ for the power approximation; $t_{40} = 0.90$, $p = 0.37$ for the exponential approximation; two-tailed). The measure of elevation in the loss domain exceeds its gain domain counterpart significantly ($t_{40} = 2.31$, $p = 0.01$ for the linear interpolation; $t_{40} = 1.91$, $p = 0.03$ for the power approximation; $t_{40} = 1.78$, $p = 0.04$ for the exponential approximation; one-tailed).

It must be mentioned that the picture is somewhat less clear at the level of individual subjects with 26 out of 41 participants satisfying $\delta^- > \delta^+$ (for linear interpolation, power approximation, and exponential approximation), $p = 0.06$ for a sign test (one-tailed). The general conclusion is strengthened again by the results of the linear probability weighting function where the measure of elevation is higher in the loss domain for 28 out of 41 subjects (for linear interpolation, power approximation, and exponential approximation), $p = 0.01$ for a sign test (one-tailed).

The parametric estimates resulting from the linear in log-odds specification can also be used to test the duality of the probability weighting function in the framework of the two-stage model. As pointed out by Abdellaoui (2000), the duality condition for the probability weighting function (i.e., $w^-(p) = 1 - w^+(1 - p)$ for all $p$), applied to the linear-in-log-odds specification, implies $\gamma^+ = \gamma^-$ and $\delta^+ = 1/\delta^-$. The results of the paired $t$ tests of the equality $\gamma^+ = \gamma^-$ have already been presented above. They are confirmed by sign tests ($p = 0.76$ for linear interpolation and power approximation; $p = 0.35$ for the exponential approximation; two-tailed). Sign tests of the restriction $\delta^+ = 1/\delta^-$ lead to qualitatively similar results ($p = 0.21$ for linear interpolation and power approximation; $p = 0.12$ for the exponential approximation; two-tailed). In summary, a violation of the duality condition for the probability weighting function cannot be established at conventional levels of significance.

## 6. Discussion and Conclusion

This paper provides a parameter-free and a fully choice-based elicitation and decomposition of decision weights under CPT. We found that SEU is violated in a systematic fashion in both the gains and loss domains. The elicited weighting functions and the choice-based probabilities seem to be consistent with the psychological principle of diminishing sensitivity, stipulating a decrease in marginal effect as a distance from a reference point increases. The reference points are 0 and 1 for $w^+(.)$ and $w^-(.)$, and $\varnothing$ and $S$ for $W^+(.)$, $W^-(.)$ and $\hat{q}(.)$. This suggests that the subjective treatment of uncertainty, in the presence of exogenously given probabilities as well as in their absence, is mainly governed by a simple psychological principle. On this point our paper extends the previous experimental findings by Tversky and Fox (1995), Fox and Tversky (1998), Wu and Gonzalez (1999) and Kilka and Weber (2001). The similarity of the properties of judged probabilities and choice-based probabilities comes as good news for the link between the psychological concept of judged probabilities and the more standard economic concept of choice-based probabilities.

The other findings regard the shape of the utility function in both gain and loss domains and the usefulness of the introduction by CPT of a specific weighting function for losses. The paper reports the results of an experimental elicitation of utility using the trade-off method with unknown probabilities. For gains, concavity is the dominating shape of the utility function. Our results are particularly similar to those obtained under risk and without taking into account the null monetary outcome as a reference point (Wakker and Deneffe 1996, Bleichrodt and Pinto 2000, Abdellaoui 2000). For losses, no clear evidence in favor of convexity was observed, and this result is also consistent with previous findings under risk (Etchart 2003a, b).

Decision weights for the loss domain exhibit more elevation, particularly for likely events. When the two-stage model is assumed, the resulting probability weighting functions exhibit more elevation for losses. Furthermore, the hypothesis of equal curvature across domains is not rejected. This is consistent with similar findings under risk (Abdellaoui 2000). The duality condition is not contradicted by our data, suggesting that CEU might approximate CPT in particular choice situations.

One possible extension of our work is to study the robustness of the Wakker's decomposition by decomposing decision weights for different sources of uncertainty (as in Kilka and Weber 2001). An elicitation of the belief component $\hat{q}(.)$ for gains and losses is also desirable.

## References


Abdellaoui, M. (2000). Parameter-Free Elicitation of Utility and Probability Weighting Functions, Management Science 46, 1497-1512.

Bleichrodt, H., J. L. Pinto (2000). A Parameter-Free Elicitation of the Probability Weighting Function in Medical Decision Analysis, Management Science 46, 1485-1496.

Ellsberg, D. (1961). Risk, Ambiguity, and the Savage Axioms, Quarterly Journal of Economics 75, 643-669.

Epstein, L. G. (1999). A Definition of Uncertainty Aversion, Review of Economic Studies 66, 579-608.

Etchart, N. (2003a). Traitement subjectif du risque et comportement individuel devant les pertes: une étude expérimentale, unpublished PhD dissertation, GRID- ENSAM, Paris.

Etchart, N. (2003b). Is Probability Weighting Sensitive to the Magnitude of Consequences? An Experimental Investigation on Losses, Journal of Risk and Uncertainty (forthcoming).

Fellner, W. (1961). Distortion of Subjective Probabilities as a Reaction to Uncertainty, Quarterly Journal of Economics 75, 670-689.

Fennema, H., M. van Assen (1999). Measuring the Utility of Losses by Means of the Tradeoff Method, Journal of Risk and Uncertainty 17, 277-295.

Fox, C. R., B. A. Rogers, A. Tversky (1996). Option Traders Exhibit Subadditive Decision Weights, Journal of Risk and Uncertainty 13, 5-17.

Fox, C. R., K. E. See (2003). Belief and Preference in Decision under Uncertainty, in Thinking: Psychological Perspectives on Reasoning, Judgment and Decision Making, D. Hardmann and L. Marcchi eds, John Wiley.

Fox, C. R., A. Tversky (1998). A Belief-Based Account of Decision Under Uncertainty, Management Science 44, 879-895.

Gilboa, I. (1987). Expected Utility with Purely Subjective Non-Additive Probabilities, Journal of Mathematical Economics 16, 65-88.

Gilboa, I., D. Schmeidler (1989). Maxmin Expected Utility with a Non-Unique Prior, Journal of Mathematical Economics 18, 141-153.

Goldstein, W. M., H. J. Einhorn (1987). Expression theory and the preference reversal phenomenon, Psychological Review 94, 236-254.

Gonzalez, R., G. Wu (1999). On the Shape of the Probability Weighting Function, Cognitive Psychology 38, 129-166.

Gul, F. (1991). A Theory of Disappointment Aversion, Econometrica 59, 667-686.

Hey, J. D., C. Orme (1994). Investigating Generalizations of Expected Utility Theory Using Experimental Data, Econometrica 62, 1291-1326.

Karmarkar, U. S. (1978). Subjective Weighted Utility: A Descriptive Extension of the Expected Utility Model, Organizational Behavior and Human Performance 21, 61-72.

Kilka, M., M. Weber (2001). What Determines the Shape of the Probability Weighting Function Under Uncertainty, Management Science 47, 1712-1726.

Lattimore, P. K., J. R. Baker, A. D. Witte (1992). The influence of probability on risky choice: A parametric examination, Journal of Economic Behavior and Organization 17, 377-400.

Lopes, L. L., G. C. Oden (1999). The Role of Aspiration Level in Risky Choice: A Comparison of Cumulative Prospect Theory and SP/A Theory, Journal of Mathematical Psychology 43, 286-313.

Prelec, D. (1998). The probability weighting function, Econometrica 66, 497-527.

Rabin, M. (2000). Risk Aversion and Expected Utility Thory: A Calibration Theorem, Econometrica 68, 5, 1281-1292.

Savage, L. (1954). The Foundations of Statistics, Wiley, New York.

Schmeidler, D. (1989). Subjective probability and expected utility without additivity, Econometrica 57, 571-587.

Stevens, S. S. (1957). On the Psychophysical Power Law, Psychological Review 64, 153-181.

Tversky, A. (1967). Utility theory and additivity analysis of risky choices, Journal of Experimental Psychology 75, 27-36.

Tversky, A., C. R. Fox (1995). Weighing risk and uncertainty, Psychological Review 102, 269-283.

Tversky, A., D. Kahneman (1992). Advances in Prospect Theory: Cumulative Representation of Uncertainty, Journal of Risk and Uncertainty 5, 297-323.

Tversky, A., D. J. Koehler (1994). Support Theory: A Nonextensional Representation of Subjective Probability, Psychological Review 101, 547-567.

Tversky, A., P. P. Wakker (1995). Risk attitudes and decision weights, Econometrica 63, 1255-1280.

Wakker, P. P. (2001). Testing and characterizing properties of nonadditive measures through violations of the sure-thing principle, Econometrica 69, 1039-1059.

Wakker, P. P. (2004). On the Composition of Risk Preference and Belief, Psychological Review 111, 236-241.

Wakker, P. P., D. Deneffe (1996). Eliciting von Neumann-Morgenstern Utilities When Probabilities Are Distorted or Unknown, Management Science 42, 1131-1150.

Wu, G., R. Gonzalez (1996). Curvature of the Probability Weighting Function, Management Science 42, 1676-1690.

Wu, G., R. Gonzalez (1999). Nonlinear Decision Weights in Choice under Uncertainty, Management Science 45, 74-85.

Wu, George, Jiao Zhang, and Richard Gonzalez (2003). "Decison under risk" To appear in The Blackwell Handbook of Judgment and Decision Making, Derek Koehler and Nigel Harvey (eds).

# Appendix

Figure 1: Screenshot of a typical choice task