# Essays on the measurement sensitivity of risk aversion and causal effects in education

Booij, A.S.

# 3 A Parameter-Free Analysis of the Utility of Money for the General Population under Prospect Theory*

## 3.1 Introduction

Expected utility is the reigning economic theory of rational decision making under risk. In this classical framework, outcomes are transformed by a strictly increasing utility function and prospects are evaluated by the probability-weighted average utility. Therefore, risk attitudes are solely explained by utility curvature under expected utility. For example, *risk aversion* (preferring the expected value of a prospect to the prospect itself) holds if and only if the utility function is concave, implying diminishing marginal utility. However, several decades of extensive experimentation has convincingly shown that "risk aversion is more than the psychophysics of money" (Lopes 1987): numerous studies have systematically falsified expected utility as a descriptive theory of decision making (Allais 1953; Kahneman and Tversky 1979). This descriptive inadequacy has been the main inspiration for the development of many alternative theories of individual decision making under risk (Starmer 2000). The most prominent of these non-expected utility models is prospect theory (Kahneman and Tversky 1979; Tversky and Kahneman 1992).

Prospect theory entails that besides the transformation of outcomes, probabilities are transformed by a subjective probability weighting function, reflecting sensitivity towards probabilities. Additionally, prospect theory entails that outcomes are evaluated relative to a reference point, reflecting sensitivity towards whether outcomes are better or worse than the status quo. According to prospect theory, risk attitudes are thus determined by a

* This chapter is based on Booij and van de Kuilen (2007).

combination of utility curvature,[20] subjective probability weighting, and the steepness of the utility function for negative outcomes (*losses*) relative to the steepness of the utility function for positive outcomes (*gains*), i.e. *loss aversion*. Consequently, in the prospect theory framework, the one-to-one relationship between utility curvature and risk attitudes no longer holds and the validity of classical measurements of risk attitudes can therefore be questioned, which explains why these measurements have often led to preference inconsistencies (Hershey and Schoemaker 1985) or theoretical implausibilities (Rabin 2000a).

A detailed separation of risk attitudes into a subjective probability weighting-, a utility curvature-, and a loss aversion component requires the collection of considerable data. The few studies that have used prospect theory to analyze risk attitudes for the general population were forced to adopt stringent parametric assumptions because not enough variation in the data was available for such a detailed separation of risk attitudes (e.g. Donkers et al. 2001; Tu 2005). Moreover, these studies assume homogeneity in preferences or model heterogeneity in an inflexible parametric way, such as through normally distributed random coefficients. The experimental approaches that use non-parametric methods to elicit utilities at the individual level (e.g. Abdellaoui 2000; Bleichrodt and Pinto 2000; Abdellaoui et al. 2007b) circumvent these problems, but the external validity of the results of these studies can be questioned if risk attitudes are related to socio-demographic characteristics that differ between the population and the student subject pools used in these experiments.

This chapter presents the results of an experiment that completely measures the utility part of risk attitudes for positive and negative monetary outcomes, using a large and representative sample of N = 1935 respondents from the Dutch population, in a completely parameter-free way. The utility measurement technique we use is the tradeoff method introduced by Wakker and Deneffe (1996). This method is robust against subjective probability distortion and is parameter-free in the sense that no prior assumptions about the true underlying functional form of the utility function have to be made. In addition, we obtain parameter-free measurements of "the core idea of prospect theory" (Kahneman, 2003,

---

[20] In this thesis we will refer to the outcome transformation as utility (Bleichrodt et al. 2001), though it is often designated by value function in the prospect theory literature to illustrate that the outcome transformation is a descriptive concept that may be distinct from intrinsic utility. Abdellaoui et al. (2007a) provide evidence that (for gains) the outcome transformation elicited under prospect theory coincides with strength of preference statements, and hence can be viewed as reflecting the intrinsic utility of money.

p.1457), i.e. loss aversion. Empirical research has shown that loss aversion is a major component of risk attitudes (Kahneman et al. 1991), and it can explain a wide variety of anomalous behavior in the field (Camerer 2000). Unlike previous measurements of risk attitudes, this study thus provides the first measurement of risk attitudes of the general population that is valid under (cumulative) prospect theory, does not depend on a priori assumptions about the underlying functional form of the utility function, is externally valid, and does not rule out heterogeneity of individual preferences. The dataset also allows us to test whether utility curvature is more or less pronounced for losses than for gains, whether scaling up monetary outcomes leads to a higher or a lower degree of utility curvature or loss aversion, and to relate the obtained measurements of utility curvature and loss aversion to socio-demographic characteristics.

First, the results show that utility is concave for gains and convex for losses, reflecting diminishing sensitivity as predicted by prospect theory but contradicting the classical prediction of universal concavity. In addition, our result support Rabin's (2000b, p.202) claim that diminishing marginal utility is an "implausible explanation for appreciable risk aversion, except when the stakes are very large"; we confirm that utility curvature is less pronounced than suggested by studies that, erroneously, assume the classical expected utility model. Second, our results confirm the common finding that females are more risk averse than males (e.g. Hartog et al. 2002; Cohen and Einav 2007; Fellner and Maciejovsky 2007). Unlike previous studies that were not able to unambiguously decompose the different components of risk attitudes and often ascribe this gender difference solely to differences in the degree of utility curvature, our results show that this phenomenon is to a large extent driven by the fact that females are more loss averse than males. Our results also show that highly educated respondents are less loss averse, suggesting that measurements of loss aversion based on highly educated student samples, as often used in the laboratory, lead to an underestimation of loss aversion. Finally, we did not find significant evidence for the hypothesis that both the degree of utility curvature and the degree of loss aversion increases with the size of the outcomes involved.

The remainder of this chapter is organized as follows. Section 3.2 briefly provides background information about the ongoing debate in the literature about the proper shape of the utility function. Section 3.3 discusses prospect theory, and section 3.4 provides an explanation of the measurement techniques we used to obtain parameter-free measurements of utility curvature and loss aversion at the individual level. Section 3.5 presents the experimental method, followed by the presentation of the results of the experiment in section 3.6. Section 3.7 contains a discussion of the experimental method and the main results, and section 3.8 concludes. Finally, section 3.9 presents details of the experimental instructions, and a table with unweighted results can be found in the appendix to this chapter in section 3.9.

## 3.2 Background

The separation of risk attitudes into a utility-, a loss aversion- and a probability weighting-component is crucial from a policy perspective because choice behaviour based on diminishing marginal utility is considered to be rational by most economists in the sense that these choices satisfy the fundamental axioms of expected utility (in particular the independence axiom) whereas choice behaviour based on subjective probability weighting for example, does not agree with these normatively compelling axioms (Bleichrodt et al. 2001; Wakker 2005). For instance, consider the so-called Equal Sacrifice Principle first put forth by Mill (1848). According to this principle, tax rates should be set in such a way that all people paying tax lose the same amount of utility. Hence, information about the utility that people derive from money obtained in isolation, i.e. obtained in an environment where confounding effects such as probability weighting did not play a role, is necessary. This study provides this information on the basis of a large representative dataset.

Parametric studies that provide measurements and decompositions of risk attitudes into a utility- and a probability weighting-component are numerous and, consequently, there is an ongoing debate about the shape of the utility function. For gains, most studies have corroborated that the utility function is concave-shaped, reflecting the natural intuition that each new Euro brings less utility than the Euro before and implying diminishing marginal utility (Wakker and Deneffe 1996). The debate regarding the shape of the utility function for losses has not been settled, however.

First of all, there is no consensus at present about the fundamental question whether the utility function for losses is either convex- or concave-shaped. Although the majority of studies have found a convex utility function for losses (Currim and Sarin 1989; Tversky and Kahneman 1992; Abdellaoui 2000; Etchart-Vincent 2004; Abdellaoui et al. 2007b), some studies have found the opposite result (Davidson et al. 1957; Laury and Holt 2007 (for real incentives only), Abdellaoui et al. 2008). A second point of debate is whether utility curvature for losses is either more or less pronounced compared to utility curvature for gains. More pronounced convexity for losses than for gains was found by Fishburn and Kochenberger (1979), virtually no difference in the degree of utility curvature was found by Abdellaoui et al. (2007b), and Schunk and Betsch (2006), whereas less pronounced convexity for losses than concavity for gains has been reported by Fennema and van Assen (1999), Wakker et al. (2007), and Abdellaoui et al. (2005). Finally, there is no consensus on whether utility curvature is more (or less) pronounced for larger outcomes. Increasing relative risk aversion has been found, for example, by Kachelmeier and Shehata (1992), Holt and Laury (2002, 2005), and Harrison et al. (2005a), whereas the opposite result, i.e. a decreasing relative risk aversion coefficient, has been found, for example, by Friend and Blume (1979), and Blake (1996).

There are four possible confounding factors in the aforementioned studies that are not present in the current study, and that may explain the seemingly contradictory findings. First of all, some studies assume expected utility and, thus, ignore the important role of probability weighting in risk attitudes. Second, the functional form of the utility (and probability weighting-) function are sometimes assumed beforehand and, therefore, the estimations depend critically on the appropriateness of the assumed functional form: conclusions drawn on the basis of the parameter estimates need no longer be valid if the true functional form differs from the assumed functional form. Third, most of these studies use aggregate data to estimate the different assumed functional forms, ruling out heterogeneity of individual preferences. Finally, student populations are commonly used as subjects, making the external validity of the results questionable.

## 3.3 Prospect Theory

We consider decision under risk, with $\mathbb{R}$ the set of possible monetary outcomes of gains and losses with respect to some wealth level or *reference point*. The reference point is assumed to be the status quo, i.e. the current wealth level. A *prospect* is a finite probability distribution over (monetary) outcomes. Thus, a prospect yielding outcome $x_i$ with probability $p_i$ ($i$ = 1,…,$n$) is denoted by ($p_1$:$x_1$,…, $p_n$:$x_n$). A two-outcome prospect ($p$:$x$, 1−$p$:$y$) is denoted by ($p$:$x$,

$y$) and the unit of payment for outcomes is one Euro. In this chapter and the following, *prospect theory* refers to the modern (cumulative) version of prospect theory introduced by Tversky and Kahneman (1992), that corrected the original '79 version for violations of stochastic dominance and, more importantly, can also deal with uncertainty, i.e. the case of unknown probabilities. Prospect theory entails that the value of a prospect with outcomes $x_1 \leq \ldots \leq x_k \leq 0 \leq x_{k+1} \leq \ldots \leq x_n$ is given by:

$$\sum_{i=1}^{k} \pi_i^- \mathrm{U}(x_i) + \sum_{j=k+1}^{n} \pi_j^+ \mathrm{U}(x_j). \tag{3.3.1}$$

Here U: $\mathbb{R} \to \mathbb{R}$ is a continuous and strictly increasing *utility function* satisfying U(0) = 0, and $\pi^+$ and $\pi^-$ are the *decision weights*, for gains and losses respectively, defined by

$$\pi_i^- = w^-(p_1 + \ldots + p_i) - w^-(p_1 + \ldots + p_{i-1}) \qquad \text{for } i \leq k \text{, and}$$

$$\pi_j^+ = w^+(p_j + \ldots + p_n) - w^+(p_{j+1} + \ldots + p_n) \qquad \text{for } j > k. \tag{3.3.2}$$

Here $w^+$ is the *probability weighting function for gains* and $w^-$ is the *probability weighting function for losses*, satisfying $w^+(0) = w^-(0) = 0$ and $w^+(1) = w^-(1) = 1$, and both strictly increasing and continuous. Thus, the decision weight of a positive outcome $x_i$ is the marginal $w^+$ contribution of $p_i$ to the probability of receiving better outcomes, and the decision weight of a negative outcome $x_i$ is the marginal $w^-$ contribution of $p_i$ to the probability of receiving worse outcomes. Finally note that the decision weights do not necessarily sum to 1 and that prospect theory coincides with expected utility if people use a fixed reference point and do not distort probabilities, i.e. prospect theory coincides with expected utility if individuals use a fixed reference point in terms of wealth *and* $w^+$ and $w^-$ are the identity.

## 3.4 Measuring The Utility Function

This section provides an explanation of the measurement techniques used to obtain parameter-free measurements of utility curvature and loss aversion at the individual level.

### *3.4.1 Measuring Utility Curvature: The Tradeoff Method*

The (gamble-) tradeoff method, introduced by Wakker and Deneffe (1996), draws inferences from a series of indifferences between two-outcome prospects in order to obtain a so-called *standard sequence of outcomes*, i.e. a series of outcomes that is equally spaced in utility units.

Contrary to other elicitation techniques often used to measure individual utility functions such as the certainty equivalence method, the probability equivalence method, and the lottery equivalence method (McCord and de Neufville 1986), utilities obtained through the tradeoff method are robust to subjective probability distortion. Hence, besides being valid under expected utility, the tradeoff method retains validity under prospect theory, rank-dependent utility and cumulative prospect theory (Wakker and Deneffe 1996).

Consider an individual who is indifferent between the prospects ($p$:$x_1$, $g$) and ($p$:$x_0$, $G$) with $0 \leq g \leq G \leq x_0 \leq x_1$. In most existing laboratory experiments employing the tradeoff method (as well as in our experiment) individual indifference is obtained by eliciting the value of outcome $x_1$ that makes a person indifferent between these two prospects while fixing outcomes $x_0$, $G$, $g$, and probability $p$. Under prospect theory, indifference between these prospects implies that:

$$w^+(p)\big(\mathrm{U}(x_1) - \mathrm{U}(x_0)\big) = \big(1 - w^+(p)\big)\big(\mathrm{U}(G) - \mathrm{U}(g)\big). \tag{3.4.1}$$

Thus, under prospect theory, the weighted improvement in utility by obtaining outcome $G$ instead of outcome $g$ is equivalent to the weighted improvement in utility by obtaining outcome $x_1$ instead of outcome $x_0$. Now suppose that the same person is also indifferent between the prospects ($p$:$x_2$, $g$) and ($p$:$x_1$, $G$). If we apply the prospect theory formula to this indifference we find that:

$$w^+(p)\big(\mathrm{U}(x_2) - \mathrm{U}(x_1)\big) = \big(1 - w^+(p)\big)\big(\mathrm{U}(G) - \mathrm{U}(g)\big). \tag{3.4.2}$$

Combining equations (3.4.2) and (3.4.1) yields:

$$\mathrm{U}(x_2) - \mathrm{U}(x_1) = \mathrm{U}(x_1) - \mathrm{U}(x_0). \tag{3.4.3}$$

Thus, the tradeoff in utilities between receiving outcome $x_2$ instead of outcome $x_1$ is equivalent to the tradeoff in utilities between receiving outcome $x_1$ instead of outcome $x_0$. Or, put differently, $x_1$ is the utility-midpoint between outcome $x_0$ and outcome $x_2$ and the sequence of outcomes $x_0$, $x_1$, $x_2$ is equally spaced in terms of utility units. One can continue eliciting individual indifference between prospects ($p$:$x_i$, $g$) and ($p$:$x_{i-1}$, $G$) in order to obtain an increasing sequence $x_0,\ldots,x_n$ of gains that are equally spaced in utility units. A similar process can be used to construct a decreasing sequence of equally spaced losses. More specifically, individual indifference between the prospects ($p$:$y_i$, $l$) and ($p$:$y_{i-1}$, $L$) with $0 \geq l \geq L \geq y_0 \geq y_1 \geq \ldots \geq y_n$ implies that the resulting decreasing sequence of losses $y_0,\ldots, y_n$ is equally spaced in utility units. In what follows, we will use the term *utility increment* (*utility* *decrement*) to denote the equal utility difference between the elements of an increasing (decreasing) standard sequence of gains (losses).

### 3.4.2 *Measuring Loss Aversion*

The tradeoff method allows measuring utilities for either gains or losses. Without further information these measurements cannot be combined, because they are not on the same scale. This requires the elicitation of additional indifferences that involve mixed prospects, i.e. prospects that involve both gains and losses. With the proper use of the obtained standard sequences, this can be done in a parameter-free way.

A full description of the utility function over gains and losses allows for the comparison of utility in the two domains and, hence, allows for the characterization of loss aversion. Unfortunately, a commonly accepted definition of loss aversion does not exist in the literature (Abdellaoui et al. 2007b). We define the loss aversion coefficient $\lambda$ as follows:

$$\lambda = \frac{\mathrm{U}(y_0)}{\mathrm{U}(x_0)}\frac{x_0}{y_0} \;. \tag{3.4.4}$$

This definition approximates the definition proposed by Köbberling and Wakker (2005), who characterize loss aversion as the ratio between the left and right derivatives of the utility function at zero, i.e. $\lambda^{KW} \equiv \mathrm{U}'_{\uparrow}(0)/\mathrm{U}'_{\downarrow}(0)$. This "local" definition measures the kink of the utility function at the reference point. It closely resembles the definition of loss aversion of Tversky and Kahneman (1992), who implicitly used $\lambda$ = U(–\$1)/U(\$1). Köbberling and Wakker's definition, however, has the advantage that it does not depend on the unit of the outcomes, i.e. it is independent of the currency unit. Other definitions have also been proposed, such as Kahneman and Tversky's (1979) original formulation of loss aversion as $-\mathrm{U}(-x) > \mathrm{U}(x)$ for all $x > 0$, or a stronger version formulated by Wakker and Tversky (1993) given by $\mathrm{U}'(-x) > \mathrm{U}'(x)$ for all $x > 0$. These definitions do not yield an index of loss aversion but formulate it as a property of the utility function over a whole range. An index can then be constructed by taking the mean or median value of the relevant values of $x$. This is not an arbitrary choice, however, making comparison between measurements difficult. Hence, we have to be careful with comparing loss aversion estimates (see Abdellaoui et al. 2007b for a more extensive discussion).

Our method to measure loss aversion consists of three steps. First we determine the utility of $x_0$ in terms of utility increments. To do so, we elicit the value of outcome $b$ that makes an agent indifferent between the prospects ($r$:$b$, 0) and ($r$:$x_1$, $x_0$), where $r$ is some fixed probability. Under prospect theory, indifference between these prospects implies:

$$\mathrm{U}(x_0) = \frac{w^+(r)}{1-w^+(r)}\big(\mathrm{U}(b) - \mathrm{U}(x_1)\big). \tag{3.4.5}$$

For given probability weights this equation pins down the utility of outcome $x_0$ in terms of increments of the standard sequence for gains. Suppose for example that $w^+(r) = 1 - w^+(r)$, and that $b$ falls within the standard sequence for gains, say half-way between $x_2$ and $x_3$. Then, using linear interpolation to obtain the utility of $b$, equation (3.4.5) implies that the utility of $x_0$ is equal to 1.5 times the utility increment for gains, i.e. $\mathrm{U}(x_0) = 1.5\big(\mathrm{U}(x_1) - \mathrm{U}(x_0)\big)$. The linear approximation of the utility of outcome $b$ can be justified on the grounds that utility is often found to be linear over small monetary intervals (Wakker and Deneffe 1996).[21] Graphically, equation (3.4.5) identifies the utility of outcome $x_0$ in terms of the amount of steps of the standard sequence, as illustrated by brace 1 in Figure 3.1.

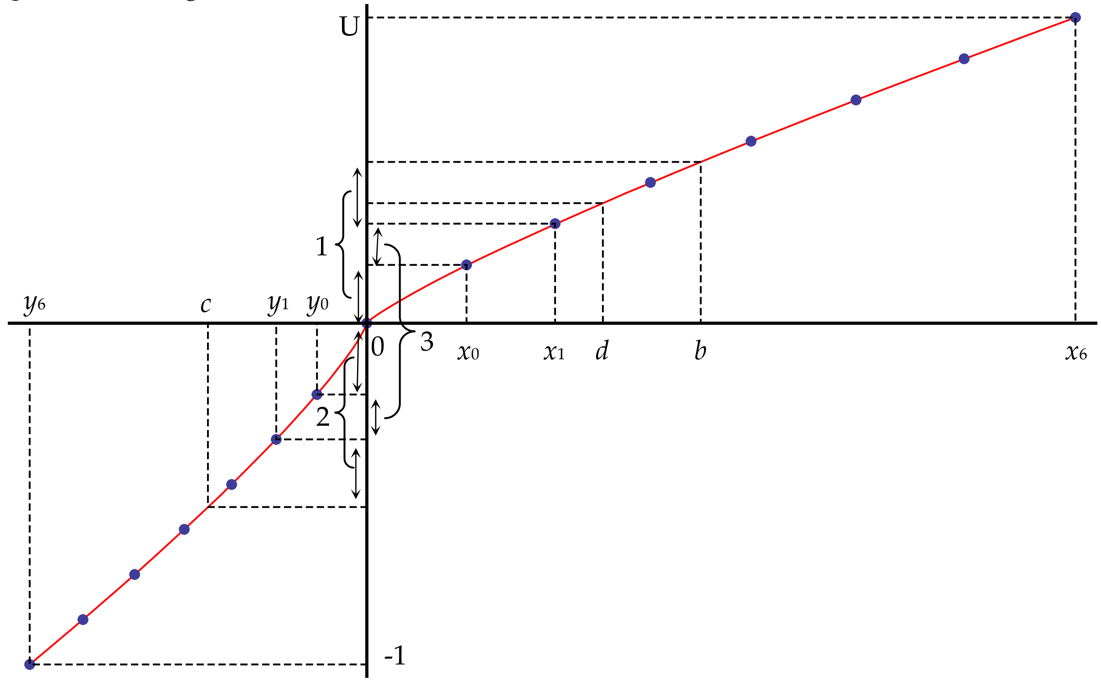
In the second step, an analogous question for losses allows for the identification of the utility of outcome $y_0$ in terms of utility decrements. We obtain the outcome $c$ that makes an agent indifferent between the prospects ($r$:0, $c$) and ($r$:$y_0$, $y_1$). Under prospect theory, indifference between these prospects implies:

$$\mathrm{U}(y_0) = \frac{w^-(1-r)}{1-w^-(1-r)}\big(\mathrm{U}(c) - \mathrm{U}(y_1)\big). \tag{3.4.6}$$

In similar spirit to equation (3.4.5) for gains, this equation identifies the utility of outcome $y_0$ in terms of utility decrements of the standard sequence for losses, which is illustrated by brace 2 in Figure 3.1 below. The outcome $c$ must fall within the standard sequence for losses and its utility can be obtained by using (linear) interpolation again.

[21] Linear interpolation to obtain the utility of outcome $b$ is only possible if outcome $b$ falls within the standard sequence, which is why we used $x_0$ and $x_1$ as outcomes for the second prospect.

**Figure 3.1:** Linking the Utilities for Gains and Losses




The first two steps connect the standard sequence for gains and losses to the zero outcome, but do not determine their relative scale. In the third and final step, the size of the utility increment for gains in terms of the utility decrement for losses is determined by eliciting the outcome $d > x_0$ that makes the agent indifferent between the mixed prospects $(r{:}d, y_1)$ and $(r{:}x_0, y_0)$. Under prospect theory, indifference between these prospects implies:

$$\mathrm{U}(y_0) - \mathrm{U}(y_1) = \frac{w^+(r)}{w^-(1-r)}\big(\mathrm{U}(d) - \mathrm{U}(x_0)\big). \tag{3.4.7}$$

From a measurement perspective this equation amounts to relating the utility decrement of the standard sequence of losses to the utility increment of the standard sequence of gains, as depicted by brace 3 in Figure 3.1. The utility of outcome $d$ can again be approximated by using interpolation.

Equations (3.4.5) – (3.4.7), the fact that the standard sequences are equally spaced in utility units, and (linear) interpolation of the utility of indifference outcomes $b$, $c$ and $d$, determines the utilities of the outcomes $\{y_n, \ldots, y_0, 0, x_0, \ldots, x_n\}$ up to scale. Setting, for example, $U(y_n) = -1$ completely pins-down the utility function such that it can be depicted graphically.[22]

In the above steps, the probability weights corresponding to the probabilities used in the elicitation procedure were assumed to be known, while in fact they are unknown a priori. Several parameter-free techniques to obtain these probability weights have been proposed in the literature (Abdellaoui 2000; Bleichrodt and Pinto 2000). Hence, if combined with these measurement methods, the three indifferences stated above can in principle be used to measure the utilities of the standard sequences of gains and losses on the same scale. In this chapter, we do not use additional questions to obtain the probability weights, and assume either linear probability weighting as in classical economic analyses or employ the empirical estimates of the probability weights found by Kahneman and Tversky (1992) in the analysis. A different method to measure loss aversion in a parameter-free way is in Abdellaoui et al. (2007b).

## 3.5 The Experiment: Method

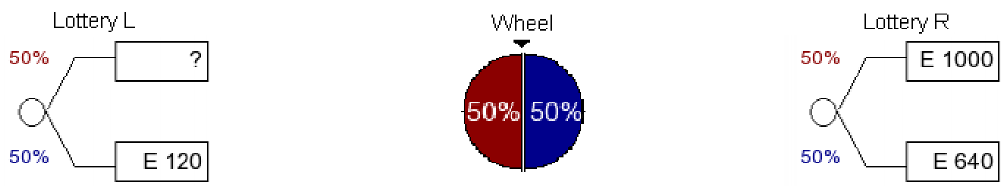
*Participants*. N = 1935 Dutch participated in the experiment which was held in February 2006. We used the DNB Household survey which is a household panel that completes a questionnaire every week on the Internet or, if Internet is not available in the household, by a special box connected to the television. The household panel is a representative sample of the Dutch population; it consists of people from all income groups and age groups above 16 years (see second column of Table 3.2 for further details).

*Procedure*. The experiment consisted of two parts, I and II. The questions of the first part (Q1 – Q16) were designed to elicit utility, as described in section 3.4. The questions of the second part (Q17 – Q27) were designed to determine probability weighting for both the domain of gains and the domain of losses. A description and an analysis of these question is given in

[22] The specific utilities for this normalization are $U(y_i) = (i-n)\Delta U_y - 1$ and $U(x_i) = i\Delta U_x + U(x_0)$ for $i = 0,\ldots,n$, with $\Delta U_y = -(1-w^-(r))/(w^-(r)(\hat{c}-n)+n)$, $\Delta U_x = (\Delta U_y w^-(r))/(w^+(r)\hat{d})$ and $U(x_0) = (w^+(r)\hat{b}\Delta U_x)/(1-w^+(r))$. The hat variables are continuous quantities that measure the utility distance between the corresponding outcome and the same ranked outcome of the other prospect, in terms of the amount of utility steps of the standard sequence of that domain. Under linear interpolation for example, we have $\hat{b} = (b-x_k)/(x_{k+1}-x_k)+k-1$ with $x_k < b < x_{k+1}$.

chapter 4. In this chapter only the questions of the first part are used. Respondents first read experimental instructions for part I (see section 3.9.2) and were then asked to answer a practice question to familiarize them with the experimental setting. In the instructions it was emphasized that there were no right or wrong answers. In order to obtain indifference between prospects we used direct matching, that is, respondents were asked to report an outcome of a prospect for which they would be indifferent between two particular prospects, that were framed as depicted in Figure 3.2 below.

**Figure 3.2:** The Framing of the Prospect Pairs in Part I




Respondents were thus simply asked to report the upper prize of the left prospect that would make them indifferent between both prospects. The wheel in the middle served to explain probabilities to respondents. Both the probabilities reported in the wheel and the colors of the wheel corresponded to the probabilities of the prospects. The prizes of the prospects used were hypothetical (for a discussion see section 3.7).

**Table 3.1:** The Obtained Indifferences in Part I

| Matching Question | Prospect L | | Prospect R |
|---|---|---|---|
| 1 | $(0.5: \underline{a}, 10)$ | ~ | $(0.5: 50, 20)$ |
| 2 | $(0.5: \underline{x_1}, g)$ | ~ | $(0.5: x_0, G)$ |
| ⋮ | ⋮ | ⋮ | ⋮ |
| 7 | $(0.5: \underline{x_6}, g)$ | ~ | $(0.5: x_5, G)$ |
| 8 | $(0.5: \underline{y_1}, l)$ | ~ | $(0.5: y_0, L)$ |
| ⋮ | ⋮ | ⋮ | ⋮ |
| 13 | $(0.5: \underline{y_6}, l)$ | ~ | $(0.5: y_5, L)$ |
| 14* | $(0.5: \underline{b}, 0)$ | ~ | $(0.5: x_1, x_0)$ |
| 15* | $(0.5: 0, \underline{c})$ | ~ | $(0.5: y_0, y_1)$ |
| 16* | $(0.5: \underline{d}, y_1)$ | ~ | $(0.5: x_0, y_0)$ |

*Notes*: underlined outcomes are the matching outcomes and questions marked with an asterisk were presented in randomized order.

*Stimuli*. For each respondent we obtained a total of 16 indifferences; see Table 3.1. Following the first practice question, matching questions 2 to 7 served to obtain an increasing sequence of gains $x_0,\ldots, x_6$ that are equally spaced in utility units, followed by six matching questions to obtain a decreasing sequence of losses $y_0,\ldots, y_6$ that are equally spaced in terms of utility (see section 3.4.1). Matching questions Q14-Q16 served to obtain a parameter-free measurement of the degree of loss aversion at the individual level (see section 3.4.2). As can be seen in Table 3.1, the parameter values of $p$ and $r$ used throughout section 3.4 were set at 1/2, as in Bleichrodt and Pinto's (2000) experiment.

*Treatments*. In order to be able to test whether utility curvature is more pronounced for larger monetary outcomes, respondents were randomly assigned to two different treatments. These treatments only differed in the parameter values used for $G$, $g$, $x_0$, $L$, $l$, and $y_0$. In the low-stimuli treatment, these parameter values were set at $G = 64$, $g = 12$, $x_0 = 100$, $L = -32$, $l = -6$, and $y_0 = -50$. In the high-stimuli treatment, all parameter values were scaled up by a factor 10, i.e. the parameter values were set at $G = 640$, $g = 120$, $x_0 = 1000$, $L = -320$, $l = -60$, and $y_0 = -500$.

## 3.6 The Experiment: Results

This section presents both parametric and non-parametric results regarding the shape of the utility function for gains and losses (Section 3.6.2), and loss aversion (Section 3.6.4). Since the panel of respondents is a representative sample of the Dutch population, the raw sample means would provide unbiased estimates of the population means if we would observe answers from all individuals in the panel. Not surprisingly, there is non-response in the data which, if non-random, will bias the statistical inferences. We deal with this sample selection problem by constructing sampling weights. The sample selection process is described in the next subsection.

### *3.6.1 Sample Selection*

Because we did not give financial incentives to participate in the experiment, we expected that some members of the panel would not respond, or would not provide answers to all questions. Indeed, about 20% of the panel members did not start the experiment or stopped answering questions before reaching the end of the experiment. Non-response is the first of four types of sample selection that we observed in the dataset.

The second type of sample selection is the result of an imposed monotonicity condition: the obtained standard sequence of gains (losses) had to be strictly increasing (decreasing), which precludes violations of stochastic dominance. This is a natural requirement, since individuals that violate it either have a lack of understanding or make a mistake, indicating a lack of attention. A number of studies have shown that not-providing incentives not only increases the variance of subjects' responses, but often also decreases performance (Camerer and Hogarth 1999; Hertwig and Ortmann 2001). We found that in our sample about 40% of the subjects displayed at least one inconsistency, the observed probability of which is heavily correlated with education (discussed below). This effect is also found by von Gaudecker et al. (2008), and it suggests that we should be cautious of sample selection with respect to education when implementing research designs targeted to the general population that involve cognitively demanding tasks.

The final two types of sample selection involve only a small fraction of respondents. Since estimated means are sensitive to outliers, we dropped extreme responses that would solely have a significant effect on the results. Observations were defined as outliers when the subjects' answer belonged to the top or bottom percentile of the distribution of outcomes for individuals who answered all questions (including the inconsistent observations). This amounts to about 1% of the observations. Using sample medians to reduce the effect of outliers would be another way to mitigate their influence, but correcting this estimator for sample selection is not straightforward. Hence, we chose to drop a small fraction of extreme observations and focused our attention on estimating population means for the general public. The last type of sample selection involves the loss aversion questions, where we imposed the condition that the values of outcomes $b$, $c$, and $d$, had to lie in the interval of the obtained standard sequence of the individual, and, hence, dropped 12 observations. The absolute frequencies of the classification of the sample selection mechanism are reported at the bottom of Table 3.2. The samples for gains, losses and loss aversion differ because different consistency conditions have been imposed, as already mentioned.

Selection with respect to education (and other variables) may introduce a bias in the estimation if the outcome variable of interest is related to it. Then information on the selection mechanism is needed to conduct proper statistical inference. Since we have background information on both in- and out-of-sample respondents we can check whether the answers of both groups of respondents differ systematically, and construct sampling

**Table 3.2:** Sample Selection Probit and Classification

| Variable | Fraction | Gains | Losses | Loss Aversion |
|---|---|---|---|---|
| Low Amounts Treatment | 50% | 0.144** | 0.101* | 0.088 |
| | | (0.058) | (0.059) | (0.065) |
| Female | 46% | −0.052 | −0.079 | −0.116* |
| | | (0.060) | (0.061) | (0.067) |
| Lower Secondary Education | 26% | 0.127 | 0.029 | 0.218 |
| | | (0.142) | (0.143) | (0.177) |
| Higher Secondary Education | 14% | 0.335** | 0.148 | 0.507*** |
| | | (0.151) | (0.152) | (0.182) |
| Intermediate Vocational Training | 19% | 0.046 | −0.148 | 0.147 |
| | | (0.146) | (0.148) | (0.180) |
| Higher Vocational Training | 25% | 0.258* | 0.039 | 0.394** |
| | | (0.143) | (0.144) | (0.175) |
| Academic Education | 11% | 0.488*** | 0.305* | 0.681*** |
| | | (0.158) | (0.158) | (0.187) |
| Age 35-44 | 18% | −0.157* | −0.202** | −0.284*** |
| | | (0.092) | (0.093) | (0.099) |
| Age 45-54 | 22% | −0.234*** | −0.281*** | −0.325*** |
| | | (0.088) | (0.089) | (0.095) |
| Age 55-64 | 18% | −0.313*** | −0.340*** | −0.524*** |
| | | (0.094) | (0.096) | (0.106) |
| Age 65+ | 19% | −0.462*** | −0.428*** | −0.632*** |
| | | (0.095) | (0.096) | (0.108) |
| € 1.150≤Income<€ 1.800 | 25% | 0.196* | 0.269** | 0.319** |
| | | (0.118) | (0.122) | (0.146) |
| € 1.800≤Income<€ 2.600 | 31% | 0.200* | 0.253** | 0.381*** |
| | | (0.115) | (0.119) | (0.143) |
| Income≥€ 2.600 | 35% | 0.344*** | 0.411*** | 0.526*** |
| | | (0.115) | (0.119) | (0.142) |
| Catholic | 30% | 0.014 | 0.008 | −0.007 |
| | | (0.068) | (0.069) | (0.077) |
| Protestant | 20% | 0.160** | 0.126 | 0.215** |
| | | (0.077) | (0.078) | (0.084) |
| Constant | | −0.503*** | −0.510*** | −1.199*** |
| | | (0.176) | (0.179) | (0.217) |
| N | | 1935 | 1935 | 1935 |
| Non-response | | 375 | 422 | 1361† |
| Non-Monotone | | 728 | 811 | 123 |
| Outside interpolation int. | | | | 12 |
| Outlier | | 18 | 12 | 1 |
| Sample | | 814 | 690 | 438 |
| Log-likelihood | | −1276.243 | −1226.928 | −978.2592 |

*Notes*: Standard errors allow for clustering within households. */**/***: significant at the 10/5/1% level. †: Combines non-response (2 cases) with not being in the union of the sample for gains and losses (1359 cases).

weights by calculating the inverse of the probability of being observed in the sample. Table 3.2 presents the results of probit regressions of appearance in the respective sub-samples with respect to gender, age, education, income and religion. All variables have been split up in categories such that the results can not driven by assumptions regarding the functional form.

As can be seen in the table, there appears to be selection on most variables. Since the loss aversion equation compounds the selection effects of the gain and the loss domains, it shows the most prominent effects. Generally, the probability that a given respondent is observed in the sample is increasing in income and education level, and decreasing in age. Women are less likely to answer our loss aversion questions, while individuals with a protestant background have a higher inclination to respond. These effects are interesting in their own right, and they are consistent with the findings of von Gaudecker et al. (2008) who specifically investigate sample selection in the same internet panel using risk aversion (choice) tasks. They find that non-response is related to education, gender and age, and that the amount of inconsistencies is more than twice as high in the cross section of the population compared to a sub-sample of young students, i.e. the typical subject pool used in laboratory experiments. Thus, respondents from the population appear to have more difficulty conducting the risk aversion choice tasks than the typical subject in the lab. We should note that since each question poses a new test for monotonicity, a random error model would predict the exclusion of a higher fraction of observations as the amount of questions rises.

For the purpose of this analysis it is important to acknowledge that there is sample selection with respect to demographic variables that, if related to utility curvature or loss aversion, will bias the population estimates if it is not accounted for. Hence, in what follows, the reported coefficients are estimates of the population means for the general public, obtained by weighting each observation by the inverse of the (predicted) probability of being included in the sample. These coefficients are unbiased estimates of the population means for the Dutch population under the assumption that non-response is random after conditioning on the selection variables that we observe. This rules out selection on the outcome variables of interest. Harrison et al. (2007) find that reducing show-up fees increases the likelihood of attracting risk-averse agents. We conjecture that this mechanism is not at work in our setting because the panel participants are used to participating in the questionnaire without pay.

Generally, the weighting scheme slightly reduces the estimated means, but leaves the qualitative pattern unaffected. We will only comment on the weighted parameters, except when there are notable differences compared to the unweighted data.[23]

### 3.6.2 *Utility Curvature: Non-Parametric Analysis*

Table 3.3 summarizes the results regarding the obtained utility function for monetary gains and losses under the different treatments. As can be seen in the table, the difference between the successive elements of the average standard sequences is mostly increasing for both gains and losses and under both treatments. Also, at face value utility curvature seems to be more pronounced for lower monetary outcomes.

**Table 3.3:** Weighted Mean Results Utility Curvature

| | GAINS | | | | LOSSES | | | |
|---|---|---|---|---|---|---|---|---|
| | High (N = 383) | | Low (N = 431) | | High (N = 330) | | Low (N = 360) | |
| $i$ | $x_i$ | $x_i - x_{i-1}$ | $x_i$ | $x_i - x_{i-1}$ | $y_i$ | $\|y_i - y_{i-1}\|$ | $y_i$ | $\|y_i - y_{i-1}\|$ |
| 1 | 1984 | 984 | 204 | 104 | –849 | 349 | –87 | 37 |
| | (602) | (602) | (93) | (93) | (231) | (231) | (37) | (37) |
| 2 | 2975 | 991 | 318 | 114* | –1242 | 393*** | –126 | 39* |
| | (1122) | (583) | (192) | (136) | (432) | (244) | (59) | (31) |
| 3 | 4017 | 1042** | 440 | 122*** | –1663 | 420*** | –168 | 42** |
| | (1665) | (655) | (339) | (165) | (641) | (254) | (83) | (31) |
| 4 | 5093 | 1076** | 577 | 137** | –2073 | 410 | –211 | 43* |
| | (2260) | (690) | (624) | (305) | (864) | (253) | (105) | (29) |
| 5 | 6179 | 1086 | 729 | 151** | –2488 | 415 | –255 | 43 |
| | (2897) | (740) | (953) | (357) | (1078) | (255) | (129) | (30) |
| 6 | 7310 | 1131** | 892 | 164** | –2910 | 422 | –299 | 45 |
| | (3597) | (820) | (1346) | (434) | (1308) | (270) | (156) | (32) |

*Notes*: Estimated averages in euros. Standard deviations in parenthesis. */**/*** indicates that the predicted population mean is significantly higher than its predecessor at the 10/5/1% level, based on weighted t-tests.

We performed t-tests to check whether the differences between the successive elements of the standard sequence for gains and losses are indeed significantly increasing.[24] As can be seen in the table, a total of 8 differences between the obtained successive elements of the standard sequence appear to be significantly increasing for gains. In the loss domain there

[23] Table 3.8 in the Appendix to chapter 3 provides summary statistics of the unweighted data.
[24] Non-parametric tests on the raw data yield similar results.

are 5 significant differences. Only one mean difference is lower than its predecessor, but this difference is not significant. Overall, these results thus imply concave utility for gains and convex utility for losses, reflecting *diminishing sensitivity* toward outcomes: people are more sensitive to changes near the status quo than to changes remote from the status quo, as predicted by prospect theory but contrary to the classical prediction of universal concavity. This result is consistent with the findings of other parameter-free studies employing the tradeoff method to obtain utilities for gains (Wakker and Deneffe 1996; Abdellaoui 2000) and losses (Fennema and van Assen 1998; Etchart-Vincent 2004) as well as with results from studies using parametric fittings (Currim and Sarin 1989; Tversky and Kahneman 1992; Heath et al. 1999; Davies and Satchell 2003).

**Table 3.4:** Classification of Respondents

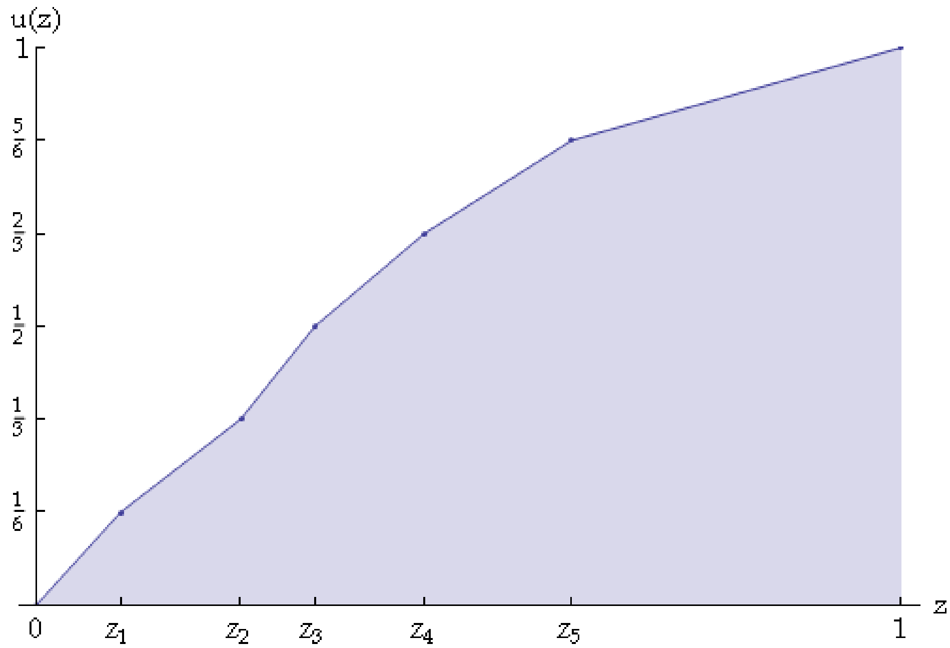
| | | LOSSES | | | |
|---|---|---|---|---|---|
| | | Concave | Linear | Convex | Total |
| GAINS | Convex | 50 (8.95%) | 31 (5.22%) | 77 (14.88%) | 158 (29.06%) |
| | Linear | 25 (4.90%) | 108 (17.54%)* | 43 (6.78%) | 176 (29.23%) |
| | Concave | 49 (8.61%) | 44 (7.68%) | 149 (25.42%)* | 242 (41.72%) |
| | Total | 124 (22.47%) | 183 (30.44%) | 269 (47.09%) | 576 (100%) |

*Note*: Table reports absolute frequencies for individuals that gave consistent answers for both gains and losses. The estimated population fractions are in parenthesis.

* A test of equality of probability is rejected in favor of the hypothesis that the concave-convex shape is more prevalent than the linear-linear form (one-sided test, z = 2.84, p-value = 0.002).

Additionally, we analyzed the shape of the utility function at the individual level. For each respondent we normalized utility of gains such that it is exactly in the unit square, i.e. we rescaled $z_i = (x_i - x_0)/x_6$ and set $u(z_i) = i/6$ for all $i = 0,\ldots,6$. Then, we classified a respondent's utility function as concave (convex; linear) if this area was larger than (smaller than; equal to) 1/2 for gains. Figure 3.3 illustrates the area measure for a typical standard sequence**.** Using a similar procedure for losses, a utility function was classified as concave (convex; linear) if the area was smaller than (larger than; equal to) 1/2. Table 3.4 gives the absolute sample frequencies and, in parenthesis, the predicted population fractions. There appears to be considerable variability in the elicited shapes of utility, with a significant amount of observations in all cells. The predominant pattern is a concave-convex shape of the utility function (25%), again implying diminishing sensitivity in both domains. If we look at unconditional probabilities, the case for diminishing sensitivity is stronger, with 42% of respondents being classified as concave for gains and 47% as convex for losses. These results are somewhat below the relative frequencies reported by Abdellaoui (2000) and Abdellaoui et al. (2007b).

**Figure 3.3:** Area under a normalized utility function




### *3.6.3 Utility Curvature: Parametric Analysis*

Although we obtained utilities in a parameter-free way using the trade-off method, we also used parametric methods to analyze the data. For each respondent, we estimated the *power utility function* with parameter $\varrho$ for both treatments. Thus, for each respondent and for each separate domain (gains and losses) we estimated:

$$U(x) = x^{\varrho} \qquad \text{for } \varrho > 0 \tag{3.6.1}$$

$$U(x) = \ln(x) \qquad \text{for } \varrho = 0 \tag{3.6.2}$$

$$U(x) = -x^{\varrho} \qquad \text{for } \varrho < 0 \tag{3.6.3}$$

by minimizing the sum of squared residuals. For losses we first transformed the outcomes to the positive domain such that the power function is well defined. The power utility function with parameter $\varrho$ is currently the most popular parametric family for fitting utility (Wakker 2008) and is also known as the family of constant relative risk aversion (CRRA) because the ratio $-xU''(x)/U'(x)$, i.e. the *index of relative risk aversion*, is constant and equal to $1 - \varrho$. We also estimated the *exponential utility function* for both gains and losses which is defined by:

$$U(x) = 1 - \exp(-\gamma\, z) \qquad \text{for } \gamma > 0 \tag{3.6.4}$$

$$U(x) = z \quad \text{for } \gamma = 0 \tag{3.6.5}$$

$$U(x) = \exp(-\gamma\, z) - 1 \quad \text{for } \gamma < 0 \tag{3.6.6}$$

where $z = (x - x_0)/(x_6 - x_0)$. This family is also known as the family of constant absolute risk aversion (CARA) because the ratio $-U''(x)/U'(x)$, i.e. the *Pratt-Arrow measure of risk aversion*, is constant and equal to $\gamma$.

Finally, we estimated the *expo-power utility function*, introduced by Abdellaoui et al. (2007a), which is a variation of the two-parameter family proposed by Saha (1993) and which is defined by:

$$U(x) = -\exp(-z^{\delta} / \delta) \quad \text{for } \delta \neq 0 \tag{3.6.7}$$

$$U(x) = -1/z \quad \text{for } \delta = 0 \tag{3.6.8}$$

where $z = x / x_6$. This particular specification allows for both concave and convex utility functions, and a subset of this specification allows for the combination of concave utility, a decreasing Pratt-Arrow measure of risk aversion ($(1 - \delta)/x + x^{\delta-1}$) and an increasing index of relative risk aversion ($1 - \delta + x^{\delta}$), which is a desirable feature because these phenomena are often found empirically (Abdellaoui et al. 2007a). As mentioned in the introduction, the one-to-one relationship between utility curvature and risk attitudes is not valid under non-expected utility models such as prospect theory and, thus, we avoid the terms index of relative risk aversion and Pratt-Arrow measure of risk aversion in what follows.

**Table 3.5:** Estimated mean utility curvature

| | UNWEIGHTED | | | WEIGHTED | | |
|---|---|---|---|---|---|---|
| Treatment | ϱ | γ | δ | ϱ | γ | δ |
| Gains, high | 0.94 | 0.14 | 1.36 | 0.95 | 0.13 | 1.37 |
| N = 378 | (0.40) | (0.64) | (0.36) | (0.41) | (0.64) | (0.37) |
| Gains, low | 0.94 | 0.19 | 1.36 | 0.94 | 0.19 | 1.36 |
| N = 428 | (0.41) | (0.86) | (0.38) | (0.41) | (0.82) | (0.38) |
| Losses, high | 0.90 | 0.17 | 1.36 | 0.92 | 0.15 | 1.37 |
| N = 326 | (0.55) | (0.66) | (0.49) | (0.59) | (0.67) | (0.51) |
| Losses, low | 0.93 | 0.19 | 1.39 | 0.93 | 0.19 | 1.38 |
| N = 356 | (0.55) | (0.76) | (0.51) | (0.55) | (0.76) | (0.51) |

*Note*: Estimated standard deviations in parenthesis.

Table 3.5 below summarizes the average optimal parameter estimates for the different parametric specifications for each specific treatment, found by minimizing the sum of squared residuals. As can be seen in the table, the (weighted) average estimate of the power coefficient ϱ for gains is 0.95 in the high- and 0.94 in the low-stimulus treatment. This result seems to be consistent with a mean estimate for gains of 0.91 found by Abdellaoui et al.

(2007b), based on a parameter-free analysis. A two-sided t-test, adjusted for the sampling weights (as are all t-tests reported in this section), does not reject the null hypothesis that the estimated mean ϱ-parameters for gains are equal across the high- and the low-stimulus treatment ($t = 0.539$, p-value = 0.593).

Analysis of the individual ϱ-parameters on the basis of one-sided t-tests does indicate that the ϱ-coefficients for gains are significantly lower than 1 in both the high- and the low-stimulus treatment (low: $t = -3.04$, p-value = 0.001; high: $t = -2.105$, p-value = 0.018), which implies a significant overall degree of diminishing marginal utility for gains. For losses, the obtained parameter estimates are 0.92 for the high-stimulus treatment and 0.93 for the low stimulus treatment respectively, and we cannot reject the hypothesis that they are equal ($t = 0.100$, p-value = 0.918). In addition, the obtained ϱ-coefficients for losses proved to be significantly lower than 1 in both the high and the low-stimulus treatment (low: $t = -2.414$, p-value = 0.008; high: $t = -2.256$, p-value = 0.012). In the parametric estimation, we thus again find that sensitivity with respect to losses diminishes, which is in line with previous findings using non-parametric data (Fennema and van Assen 1998; Abdellaoui 2000; Etchart-Vincent 2004; Abdellaoui et al. 2007b; Schunk and Betsch 2006).

Although respondents in the sample were significantly less risk averse in the gain domain (Wilcoxon rank sum test, $z = 2.159$, p-value = 0.031), we could not draw such a strong conclusion for the population ($t = 1.315$, p-value = 0.189), thus supporting the findings of Abdellaoui et al. (2007), and Schunk and Betsch (2006) who find no differences in curvature across domains. The parameter estimates of the exponential and expo-power utility functions are all highly correlated to the estimated power parameters and, hence, statistical tests based on these functional families give very similar results and will not be reported here.

The dataset allows us to relate the degree of utility curvature to socio-demographic variables. Therefore, we performed a simple (weighted) linear regression analysis with the individual estimates as dependent variables and a treatment dummy and several socio-demographic variables as independent variables.[25] The results of this regression are reported in the first three columns of Table 3.6. As can be seen in the table, the estimated coefficients

[25] We used several alternative (non-linear) functional forms, but this did not change the results.

for gains ($\varrho^+$) and losses ($\varrho^-$) appear largely idiosyncratic, with no apparent association with any of the included demographic variables except for a weak association with age. Because our method to obtain utilities is robust to subjective probability weighting and we treat gains and losses separately, our results suggest that the differences in measured risk attitudes by gender and education that are often observed stem from differences in probability weighting or loss aversion. With respect to gender, for example, our results are consistent with a recent study by Fehr-Duda et al. (2006), who do not find gender differences in the utility functions for both gains and losses based on parametric fittings and using students as subject. This result questions the validity of ascribing gender differences in risk taking behavior to differences in utility curvature, as is done in classical studies (Barsky et al. 1997; Hartog et al. 2002; Donkers et al. 2001). Other socio-demographic variables concerning profession, affinity with financial matters, being a house owner and religion were also not found to be a significant predictor of utility.

**Table 3.6:** Regression Results

| Variable | $\varrho^+$ | $\varrho^-$ | $\lambda$ |
|---|---|---|---|
| *Low Amounts* | −0.033 | 0.847 | −0.027 |
| | (0.042) | (0.599) | (0.200) |
| *Female* | −0.052 | 0.012 | 0.416** |
| | (0.039) | (0.600) | (0.211) |
| *Age* | 0.003** | −0.003 | 0.007 |
| | (0.001) | (0.011) | (0.007) |
| *High Education* | −0.044 | −0.108 | −0.497*** |
| | (0.038) | (0.644) | (0.164) |
| ln(*Income*+1) | 0.008 | 0.047 | −0.009 |
| | (0.012) | (0.084) | (0.046) |
| *Constant* | 0.830*** | −0.007 | 1.587*** |
| | (0.101) | (0.924) | (0.555) |
| N | 811 | 686 | 437 |
| $R^2$ | .012 | .004 | .042 |

*Notes*: Weighted linear regression. Standard errors allow for clustering within households. */**/***: significant at the 10/5/1% level.

### *3.6.4 Loss Aversion*

Table 3.7 presents the summary statistics and weighted means for the different indifference values of outcomes *b*, *c*, and *d*, and the resulting loss aversion coefficient.[26] Loss aversion is calculated both on the assumption of expected utility, i.e. $w(½) = ½$, and prospect theory. For the latter we use the estimated subjective probability weighting function found by Kahneman and Tversky (1992), which implies $w^{-}(½) = 0.4540$ and $w^{+}(½) = 0.4206$. Table 3.7 shows there are significant differences in the loss aversion coefficient between the weighted and unweighted samples. The weighted mean value of $\lambda$ under expected utility, denoted by $\lambda_{EU}$, is 1.79 for the low- and 1.73 for the high-amounts treatment, while the sample averages are 1.64 and 1.69 respectively. This is caused by the fact that lower educated individuals are underrepresented in our sample, while they are more loss averse on average. Under prospect theory, the weighted mean value of $\lambda$, denoted by $\lambda_{PT}$, is equal to 1.84 for the high-stimuli treatment and 1.90 for the low-stimuli treatment. These results are suggestive of decreasing loss aversion which has been found in other studies (Abdellaoui et al. 2007b; Bleichrodt and Pinto 2002 (health)), but the standard errors are too high to draw any strong conclusions, i.e. we cannot reject the hypothesis that loss aversion is the same across treatments (t = 0.321, p-value = 0.747).

**Table 3.7:** Mean Results Loss Aversion

| | UNWEIGHTED | | WEIGHTED | |
|---|---|---|---|---|
| | High<br>N = 210 | Low<br>N = 229 | High | Low |
| *b* | 4016 | 386 | 3970 | 391 |
| | (1604) | (150) | (1553) | (148) |
| *c* | –1569 | –157 | –1572 | –163 |
| | (612) | (59) | (604) | (62) |
| *d* | 1842 | 180 | 1833 | 184 |
| | (833) | (88) | (817) | (85) |
| $\lambda_{EU}$ | 1.69 | 1.64 | 1.73 | 1.79 |
| | (1.62) | (1.43) | (1.73) | (1.49) |
| $\lambda_{PT}$ | 1.79 | 1.74 | 1.84 | 1.90 |
| | (1.72) | (1.51) | (1.83) | (1.58) |

*Note*: Standard deviations in parenthesis.

[26] Given the values *b*, *c* and *d*, the loss aversion coefficient can be written as $\lambda = (1 - w^{+}(r))\hat{c}\hat{d} / (1 - w^{-}(r))\hat{b}$ (see footnote 22).

Under both assumptions the 95% confidence intervals of the loss aversion measure are [1.58, 1.94] and [1.68, 2.06] respectively. These intervals exclude the value of 1, meaning there is significant evidence of loss averse behavior for sample and also for Dutch population as a whole. Thus, generally, our result suggests that on average people weight a particular loss about 1.87 times as heavy as a corresponding gain when making decisions. The estimated loss aversion coefficient $\lambda$ is lower than the parametric estimate of $\lambda$ = 2.25 obtained by Tversky and Kahneman (1992), and the non-parametric estimate of $\lambda$ = 2.15 based on the definition of loss aversion proposed by Kahneman and Tversky (1979), found by Abdellaoui et al. (2007b). Our (weighted) mean estimate of $\lambda$ is however more consistent with a recent study by Gächter et al. (2007) who report an average (within-subject) $\lambda$ of 1.95 in a riskless setting, using WTA and WTP disparities of a miniature car, using a large sample of car buyers. The authors also measured loss aversion at the individual level using choices between mixed prospects and find that both measures of loss aversion correlate highly, which suggests that loss aversion is a constant trait that operates in different decision contexts in a similar way.

Interestingly, if we regress the obtained measurement of loss aversion on socio-demographic characteristics, we find that females are significantly more loss averse than males as the final column of Table 3.6 shows. Conditional on other covariates, females weight losses about 0.416 more heavily than males.[27] The difference in the unconditional means is about the same, with an estimated loss aversion coefficient of 2.10 for women, and 1.66 for men. In addition, higher educated individuals, defined as having a higher vocational education or better, appear to have a much lower degree of loss aversion (–0.497), which is consistent with the unconditional effects found by Schmidt and Traub (2002) and Gächter et al. (2007). This provides evidence that, contrary to classical studies that ascribed gender differences in risk taking behavior solely to differences in the degree of utility curvature (Barsky et al. 1997; Halek and Eisenhauer 2001; Hartog et al. 2002), this phenomenon is to a large extent driven by loss aversion.

When conditioning on other covariates Gächter et al. (2007) do not find any significant gender differences, but find strong effects of age, income and education. We do not find a strong effect of either age or income, which may be caused by differences in the sample compositions, the correlation between the covariates, and also elicitation context. The fact that direction of effects is the same in both studies strengthens the case that loss aversion is a personal trait, but further research into these associations is needed, especially if we want to draw conclusions on the causal effect of education and income on loss aversion and vice-versa.

## 3.7 Discussion

### *3.7.1 Discussion of method*

We used direct matching to obtain indifferences between prospects. There is evidence that using direct choice between prospects by using a bisection method (Abdellaoui 2000) or by using a multiple price list (Tversky and Kahneman 1992; Holt and Laury 2002) to obtain indifference between prospects yields more reliable results, with fewer choices that are inconsistent within subject (Bostic et al. 1990; Luce 2000). However, using such methods to obtain indifferences is fairly time consuming which was not tractable in this large-scale experiment with the general public.

We used hypothetical incentives in our experiment. There is an extensive debate in experimental methodology about whether real or hypothetical incentives yield better or more reliable data. Camerer and Hogarth (1999) and Hertwig and Ortmann (2001) provide excellent summaries of the ongoing debate. In general, real incentives do seem to reduce data variability (Camerer and Hogarth 1999) and increase risk aversion in choice (Holt and Laury 2002, 2005; Weber et al. 2004; Harrison 2006) and direct matching tasks (Kachelmeier and Shehata 1992). We did not use the incentive compatible Becker-DeGroot-Marschak (BDM) rewarding scheme to implement real incentives for the following reasons. First of all,

---

[27] It could be argued that this is the case because females and males weight probabilities differently as chapter 4 and a recent study by Fehr-Duda et al. (2006) suggest. It can be shown, however, that our estimate of loss aversion depends proportionally on $(1 - w^{+}(½))/(1 - w^{-}(½))$. This means that pessimism with respect to probabilities *increases* our estimate of loss aversion. This implies that our obtained gender difference in loss

a large part of the experiment concerned substantial losses and, hence, real incentives could not be used for ethical reasons. Second, the BDM scheme is fairly complex (Braga and Starmer 2005) and the BDM scheme is prone to irrational auction strategies (Plott and Zeiler 2005, p. 537). For example, respondents might report a higher matching outcome thinking it is a clever bargaining strategy or respondents might fail to understand that it is a dominant strategy to report their true matching outcome. Because it is important to minimize the burden on respondents in a large-scale experiment, this was another reason for not implementing real incentives. Third, there is evidence that real incentives do not affect results in relatively simple tasks (Camerer and Hogarth 1999), and we conjecture that the trade-off method is less susceptible to hypothetical bias since it measures utility curvature through the changes in the changes in outcomes, not the levels. Hence, any bias of similar magnitude operating on all the answers will leave our measurements of utility curvature unchanged. The effect on the loss aversion questions is ambiguous.[28] Fourth and finally, due to practical limitations it is virtually impossible to implement real incentives in a large-scale experiment (Donkers et al. 2001; Guiso and Paiella 2003), although Harrison et al. (2005b) did use real incentives in their study using a representative sample of 253 individuals taken from the Danish population, as did Dohmen et al. (2006) using a representative sample of 450 individuals taken from the German population.

### 3.7.2 *Discussion of the main results*

If we compare our findings to other measurements of risk attitudes using large representative datasets, we find that our estimated relative risk coefficient for gains of 0.06 (= 1 − 0.94) is relatively small. For example, Andersen et al. (2008) found a mean risk aversion coefficient of 0.74, and Barsky et al. (1997) found a mean risk tolerance (the reciprocal of the constant relative risk coefficient) of 0.24 which translates into a mean relative risk coefficient of 4.16. The smallest degree of relative risk aversion coefficient found

---

aversion would become even stronger if we allow probabilities weights to differ between the sexes. Hence, the gender coefficient can be interpreted as a lower bound.

[28] It is hard to speculate on the effect of hypothetical bias on our measurement of loss aversion. Assuming that individuals are more risk seeking in hypothetical settings (Weber et al. 2004; Harrison 2006), $c$ will be upward biased while $d$ and $b$ will be downward biased. The effect on the resulting loss aversion coefficient, which is proportional to $\hat{c}\hat{d}/\hat{b}$, is ambiguous (see footnote 22 and 26).

by Hartog et al. (2002) was 20. Clearly, the difference between these studies and the present study is that all these studies assumed expected utility and hence ignored the important role of probability weighting in the analysis. Hence, our results give empirical support to the conjecture of Rabin (2000b, p.202) being that diminishing marginal utility is an “implausible explanation for appreciable risk aversion, except when the stakes are very large”: utility curvature is less pronounced than suggested by classical utility measurements. Hence, this suggests that the phenomenon probability weighting is valid outside the laboratory, that is, the results support the external validity of subjective probability weighting.

In addition, our results show that utility for money is concave for gains and convex for losses, supporting the presence of a reflection effect (i.e. risk attitudes for gains are the mirror image of risk attitudes for losses) as predicted by prospect theory, but contradicting the classical prediction of universal concavity. Further, we find that lower educated respondents are more loss averse which suggests that measurements of loss aversion based on (highly educated) student samples (e.g. Abdellaoui et al. 2007b) lead to an underestimation of the loss aversion coefficient. Also, our results confirm the common finding that females are more risk averse than males. Contrary to classical studies that ascribed this gender difference solely to differences in the degree of utility curvature (e.g. Barsky et al. 1997; Pålsson 1996; Hartog et al. 2002), our results suggest that this finding is to a large extent caused by gender differences in the degree loss aversion. Finally, we did not find evidence that the degree of utility curvature (e.g. Holt and Laury 2002, 2005; Friend and Blume 1979) or the degree of loss aversion (Abdellaoui et al. 2007b) de- or increased with the size of the gains and losses involved.

## 3.8 Conclusion

We have obtained parameter-free measurements of the utility component as well as the loss aversion component of risk attitudes using a representative sample from the Dutch population. The results suggest that utility is concave for gains and convex for losses, implying diminishing sensitivity towards outcomes, as predicted by prospect theory. In addition, our results suggest that classical utility measurements overestimate concavity, which can be explained by the ignoring of probability weighting and loss aversion in these measurements. In addition, we have obtained parameter-free measurements of loss aversion. The results show that respondents were significantly loss averse and weighted a loss about 1.87 times as heavy as a commensurable gain. Interestingly, we have found evidence that males and higher educated persons are significantly less loss averse. The former result suggests that gender differences in risk attitudes are primarily driven by loss aversion and not by utility curvature as suggested by previous studies that assume the classical expected utility model. The latter result suggests that measurements of loss aversion based on relatively highly educated student samples lead to an underestimation of the loss aversion coefficient. Finally, we found that both the degree of utility curvature and the degree of loss aversion are not altered by scaling up the monetary outcomes involved.

## 3.9 Appendix to chapter 3

### 3.9.1 *Unweighted results*

**Table 3.8:** Sample Mean Results Utility Curvature

| | GAINS | | | | LOSSES | | | |
|---|---|---|---|---|---|---|---|---|
| | High (N = 383) | | Low (N = 431) | | High (N = 330) | | Low (N = 360) | |
| $i$ | $x_i$ | $x_i - x_{i-1}$ | $x_i$ | $x_i - x_{i-1}$ | $y_i$ | $\lvert y_i - y_{i-1} \rvert$ | $y_i$ | $\lvert y_i - y_{i-1} \rvert$ |
| 1 | 1993 | 993 | 205 | 105 | –851 | 350 | –86 | 36 |
| | (602) | (602) | (94) | (94) | (231) | (231) | (36) | (36) |
| 2 | 3000 | 1007 | 319 | 114*** | –1243 | 392*** | –126 | 40*** |
| | (1131) | (591) | (184) | (126) | (431) | (242) | (59) | (32) |
| 3 | 4060 | 1060*** | 441 | 122*** | –1664 | 421*** | –168 | 42*** |
| | (1692) | (676) | (313) | (149) | (634) | (253) | (83) | (32) |
| 4 | 5161 | 1101** | 576 | 135* | –2075 | 411 | –211 | 43** |
| | (2311) | (717) | (561) | (271) | (856) | (251) | (106) | (29) |
| 5 | 6283 | 1122** | 727 | 151*** | –2494 | 419 | –254 | 43 |
| | (2980) | (773) | (865) | (339) | (1069) | (256) | (130) | (30) |
| 6 | 7447 | 1164** | 893 | 166 | –2920 | 426* | –298 | 44 |
| | (3713) | (860) | (1244) | (424) | (1297) | (267) | (156) | (33) |

*Notes*: standard deviations in parenthesis. */**/*** significantly higher than its predecessor at the 10/5/1% level (Wilcoxon signed-rank test).

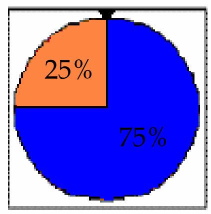
### 3.9.2 *Experimental Instructions (Part I)*

[Instructions are translated from Dutch]

Welcome at this experiment on individual decision making. The experiment is about your risk attitude. Some people like to take risks while other people like to avoid risks. The goal of this experiment is gain additional insight into the risk attitude of people living in the Netherlands. This is very important for both scientists and policymakers. If we get a better understanding of how people react to situations involving risk, policy can be adjusted to take this into account (for example with information provision on insurance and pensions, and advice for saving and investment decisions). Your cooperation at this experiment is thus very important and is highly appreciated.

The questions that will be posed to you during this experiment will not be easy. We therefore ask you to read the following explanation attentively. In this experiment, there are no right or wrong answers. It is exclusively about your own preferences. In those we are interested.

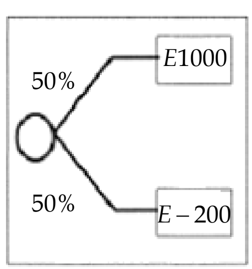
Probabilities (expressed in percentages) play an important role in this experiment. Probabilities indicate the likelihood of certain events. For example, you probably have once heard Erwin Krol say that the probability that it will rain tomorrow is equal to 20 percent




(20%). He then means, that rain will fall on 20 out of 100 similar days. During this experiment, probabilities will be illustrated using a wheel, as depicted below.

Suppose that the wheel depicted in the picture above is a wheel consisting of 100 equal parts. Possibly you have seen such a wheel before in television shows such as The Wheel of Fortune. Now imagine that 25 out of 100 parts of the wheel are orange and that 75 out of 100 parts are blue. The probability that the black indicator on the top of the wheel points at an orange part after spinning the wheel is equal to 25% in that case. Similarly, the probability that the black indicator points at an blue part after spinning the wheel is equal to 75%, because 75 out of 100 parts of the wheel are blue. The size of the area of a color on the wheel thus determines the probability that the black indicator will end on a part with that color.

Besides probabilities, lotteries play an important role in this experiment. Perhaps you have participated in a lottery such as the National Postal Code Lottery yourself before. In this experiment, lotteries yield monetary prizes with certain probabilities, similar to the

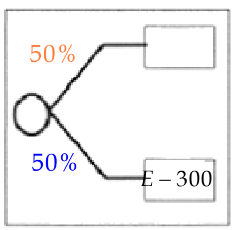


National Postal Code Lottery. However, the prizes of the lotteries in this experiment can also be negative. If a lottery yields a negative prize, you should imagine yourself that you will have to pay the about amount of money. In the following explanation we will call a negative prize a loss and a positive prize a profit. During this experiment, lotteries will be presented like the example presented below:

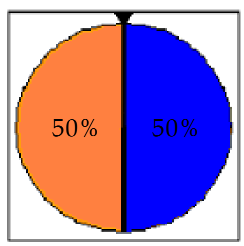
In this case, the lottery yields a profit of 1000 Euro with probability 50%. However, with probability 50%, this lottery yields a loss of 200 Euro. You should image that if you participated in this lottery, you would get 1000 Euro with probability 50%, and with probability 50% you would have to pay 200 Euro.

During this experiment you will see two lotteries, named Lottery L (Left) and Lottery R (right), on the top of each page. Between these lotteries you will see a wheel that serves as an aid to illustrate the probabilities used. You will see an example of the layout of the screen on the next page.

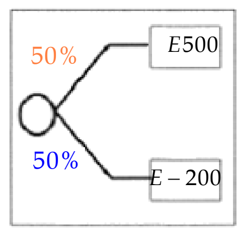
Lottery L




Wheel




Lottery R




In this example, Lottery R yields a profit of 500 Euro with probability 50% and with probability 50% it yields a loss of 200 Euro. You should imagine that, if we would spin the wheel once and the black indicator would point at the orange part of the wheel, Lottery R would yield a profit of 500 Euro. However, if the black marker would point at the blue part of the wheel, Lottery R would yield a loss of 200 Euro.

Similarly, Lottery L yields a loss of 300 Euro with probability 50%. However, as you can see, the upper prize of Lottery L is missing. During this experiment, we will repeatedly ask you for the upper prize of Lottery L (in Euro) that makes Lottery L and Lottery R equally good or bad for you. Thus, we will ask you for the upper prize of Lottery L for which you value both lotteries equally.

You could imagine that most people prefer Lottery L if the upper prize of Lottery L is very high, say 3000 Euro. However, if this prize is not so high, say 500 Euro, most people

would prefer Lottery R. Somewhere between these two prizes there is a “turnover point” for which you value both lotteries equally. For high prizes you will prefer Lottery L and for low prizes you will prefer Lottery R. The turnover point is different for everybody and is determined by your own feeling. To help you a little bit in the choice process, we will report the range of prizes in which the answer of most people lies approximately at each question.
How this works precisely will become clear in the practice question that will start if you click on the CONTINUE button below. If something it not clear to you, you can read the explanation of this experiment again by pressing the BACK button below.
[Practice question]

The practice question is now over. The questions you will encounter during this experiment are very familiar to the practice question. If you click on the BEGIN button below, the experiment will start. If you want to go through the explanation of this experiment again, click on the EXPLANATION button. Good luck.