

# A Quantitative Measurement of Regret Theory<sup>1</sup>

HAN BLEICHRODT

Department of Economics, H13-27, Erasmus University, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands.

ALESSANDRA CILLO

IESE, Avda. Pearson 21, 08034 Barcelona, Spain

ENRICO DIECIDUE

INSEAD, Boulevard de Constance, 77305 Fontainebleau Cedex, France.

March 2007

## Abstract

This paper introduces a choice-based method that for the first time makes it possible to quantitatively measure regret theory, one of the most popular models of decision under uncertainty. Our measurement is parameter-free in the sense that it requires no assumptions about the shape of the functions reflecting utility and regret. The choice of stimuli was such that event-splitting effects could not confound our results. Our findings were largely consistent with the assumptions of regret theory although some deviations were observed. These deviations can be explained by psychological heuristics and reference-dependence of preferences.

KEYWORDS: Regret theory, utility measurement, decision under uncertainty.

---

<sup>1</sup> We are grateful to Graham Loomes, Stefan Trautmann, Peter P. Wakker, George Wu, an associate editor, and two referees for their comments on previous drafts of this paper, to Alpika Mishra for writing the computer program used in the experiment, and to Arthur Attema for his help in running the experiment. Han Bleichrodt's research was supported by a VIDI grant from the Netherlands Organization for Scientific Research (NWO). Alessandra Cillo and Enrico Diecidue's research was supported by grants from the INSEAD R&D department and the INSEAD Alumni Fund (IAF).

## 1. Introduction

Regret theory, first proposed by Bell (1982) and Loomes and Sugden (1982), is one of the most popular nonexpected utility models. The theory has intuitive appeal and has a relatively simple structure relative to other nonexpected utility models. It is based on two functions: a utility function  $u(x)$ , which is defined over outcomes, and a function  $Q(u(x) - u(y))$ , defined over utility differences, which captures the impact of regret. In spite of its simple structure, regret theory can account for a wide range of observed deviations from expected utility, including the common consequence effect, the common ratio effect, the joint existence of gambling and insurance, and the preference reversal phenomenon. The crucial assumption of regret theory that generates the observed deviations from expected utility is that the function  $Q(\cdot)$  is convex. Regret theory is the only important theory of decision under uncertainty that is consistent with violations of transitivity. The violations of transitivity predicted by regret theory have been commonly observed experimentally (e.g. Loomes et al. 1991) and cannot be explained by the other main nonexpected utility theories, including prospect theory (Kahneman and Tversky 1979, Tversky and Kahneman 1992). Hence, there is clearly a domain on which it is important to study regret theory.

Many studies have tested the predictions of regret theory in a qualitative way (for an overview see Starmer 2000). They adopted the assumptions of regret theory, in particular convexity of  $Q(\cdot)$ , and then examined whether regret theory could account for their data. The validity of most of these studies was called into question by Sugden and Starmer (1993). They found that a substantial part of the empirical support for regret theory could be explained by event-splitting effects, by which an event with a given probability is weighted more heavily if it is considered as two subevents than if it is considered as a single event (see also Humphrey 1995). For psychological evidence on the impact of regret on decision making under uncertainty see, for example, Larrick (1993), Zeelenberg et al. (1996) and Zeelenberg (1999).

Several papers have analyzed the role of regret in specific domains. Smith (1996), and Yaniv (2000) analyze the role of regret in medical decision making. Gollier and Salanié (2006) and Muermann et al. (2006) incorporate regret into models of asset pricing and portfolio choice. Braun and Muermann (2004) show that regret can explain the commonly observed preference for low deductibles in personal insurance markets. Barberis et al. (2006) use regret theory to explain why people tend to invest too little in

stocks, Michenaud and Solnik (2006) show that regret can explain observed hedging behavior, and Muermann and Volkman (2007) show how regret can explain the disposition effect, the finding that investors have a preference for selling winning stocks too early and holding losing stocks too long (Shefrin and Statman 1985). The crucial assumption in most of these studies amounts to convexity of the regret function  $Q(\cdot)$ . Whether this assumption is justified is an open question as no quantitative measurements of  $Q(\cdot)$  are currently available in the literature. Providing such a quantitative measurement of  $Q(\cdot)$  is the topic of this paper.

A quantitative measurement of regret theory is not only important for descriptive purposes, but also for practical decision analysis, where prescriptions and predictions have to be made. The prescriptive status of regret theory is controversial. Several authors have argued in favor of the prescriptive validity of regret theory and intransitive choice (Loomes and Sugden 1982, Bell 1985, Anand 1987, Fishburn 1991), while others have rejected the prescriptive status of regret theory because of the possibility of intransitive preferences. Regardless of one's view on the prescriptive status of regret theory, it is important for practical decision analysis to have a method that allows quantifying regret theory. This is obvious if regret theory is considered prescriptive. However, it is also true if regret theory is merely considered descriptive. Prescriptive decision analyses require the measurement of utilities. Measuring utilities is a descriptive task and, hence, susceptible to the biases that lead to violations of expected utility. To correct for these biases a method is needed that can separate the irrational part of preferences (regret) from the rational part (utility or the attitude towards outcomes). Our measurement method makes it possible to achieve such a separation and allows measuring utility without the impact of regret.

Our measurement method is parameter-free in the sense that it requires no assumptions about the functional forms of  $u(\cdot)$  or  $Q(\cdot)$ . It is based on Wakker and Deneffe's (1996) tradeoff method. The tradeoff method was previously used to measure utility under prospect theory. We show that it can also be used to measure utility under regret theory. Consequently, regret theory can be measured by well-known and widely used measurement techniques and its measurement is much easier than commonly thought and can be based on direct choice, the basic primitive of utility theory.

We applied our method in an experiment where we controlled for event-splitting effects. We found that most subjects had a concave utility  $u(\cdot)$ ; at the aggregate level,  $u(\cdot)$  was, however, close to linear. Consistent with the predictions of regret theory,  $Q(\cdot)$  was predominantly convex both at the individual level and at the aggregate level. Some deviations existed, however, that could not be explained by regret theory and that suggest that other factors besides utility and regret affect decisions under uncertainty.

In what follows, Section 2 reviews regret theory. Section 3 describes our method for quantitatively measuring regret theory. Section 4 describes the design of our experiment and Section 5 its results. Section 6 presents predictions and implications from our results. Section 7 contains a discussion of the main findings and Section 8 concludes. Details on the experiment and an overview of the individual data are in the Appendix.

## 2. Regret Theory

Let  $\mathcal{S}$  denote a *state space*. Subsets of  $\mathcal{S}$  are called *events*. A probability measure  $p$  is given over the set of events. An *act* is a function from  $\mathcal{S}$  to the set of *outcomes*. We indicate acts as  $f, g$ . Our elicitation method, described in Section 3, requires that outcomes are real numbers. They are money amounts in the experiment in Section 4. Because our elicitation method only uses acts with two different outcomes we will restrict attention to such *binary acts*. We shall denote acts as  $(x, p; y)$ , which means that there is an event  $E$  with probability  $p$  such that  $x$  obtains under  $E$  and  $y$  obtains under the complement of  $E$ .

We assume that a preference relation  $\succsim$  is given over the set of binary acts. The conventional notation  $>$  and  $\sim$  is used to denote strict preference and indifference. By restricting attention to *constant acts*, i.e. acts for which  $x = y$ , a preference relation over outcomes can be defined, which we also denote by  $\succsim$ . We assume that higher outcomes are preferred to lower outcomes. For money amounts this assumption is self-evident.

Consider two acts  $f = (f_1, p; f_2)$  and  $g = (g_1, p; g_2)$ . The general formulation of regret theory proposed by Loomes and Sugden (1987) postulates a real-valued function  $\Psi(f_i, g_j)$  such that

$$f \succsim g \Leftrightarrow p\Psi(f_1, g_1) + (1-p)\Psi(f_2, g_2) \geq 0. \quad (1)$$

The function  $\Psi(f_j, g_j)$ ,  $j = 1, 2$ , can be interpreted as assigning a real-valued index to the net advantage of choosing  $f$  rather than  $g$  if event  $j$  obtains.  $\Psi(\cdot, \cdot)$  is unique up to scale, i.e. it can be replaced by any other function  $\Psi'(\cdot, \cdot) = \alpha\Psi(\cdot, \cdot)$ ,  $\alpha > 0$  without affecting preferences, and satisfies the following three restrictions:

- (i)  $\Psi(\cdot, \cdot)$  is *strictly increasing* in its first argument: for all outcomes  $x$  and  $y$  if  $x > y$  then  $\Psi(x, \cdot) > \Psi(y, \cdot)$ .
- (ii)  $\Psi(\cdot, \cdot)$  is *skew-symmetric*: for all outcomes  $x$  and  $y$ ,  $\Psi(x, y) = -\Psi(y, x)$ .
- (iii) for all outcomes  $x > y > z$ ,  $\Psi(x, z) > \Psi(x, y) + \Psi(y, z)$ . This property was labeled *convexity* in Loomes and Sugden (1987a), but is now generally referred to as *regret aversion*.

Skew-symmetry entails that for all outcomes  $x$ ,  $\Psi(x, x) = 0$ . Expected utility is the special case of regret theory in which  $\Psi(x, y) = u(x) - u(y)$  and  $u$  is a von Neumann Morgenstern utility function. Fishburn's (1982) skew-symmetric bilinear (SSB) theory resembles the general form of regret theory but assumes that preferences are defined over prospects, probability distributions over outcomes, rather than over acts.

Bell (1982, 1983) and Loomes and Sugden (1982) considered a restricted form of (1) in which

$$\Psi(x, y) = Q(u(x) - u(y)). \quad (2)$$

Bell (1982) refers to  $u$  in Eq. 2 as a value function measuring strength of preference, or incremental value. Loomes and Sugden (1982) refer to  $u$  as a *choiceless utility function*, which reflects the utility the decision maker would derive from an outcome  $x$  if he experienced it without having chosen it. That is, both Bell (1982) and Loomes and Sugden (1982) assume another primitive beyond choice.<sup>2</sup> Our method, which we will explain in Section 3, is entirely choice-based and we do not assume any other primitives like incremental value or choiceless utility. To emphasize this point, we will refer to  $u$  as a utility function in what follows.

The function  $Q(\cdot)$  in Eq. 2 is strictly increasing and has the following symmetry property: for all  $k$ ,  $-Q(k) = Q(-k)$ . Regret aversion, which yields the distinctive predictions of regret theory, implies that  $Q(\cdot)$  is convex. It follows from the properties of difference measurement that  $u(\cdot)$  is unique up to scale and unit

---

<sup>2</sup> In the appendix to their paper, Loomes and Sugden explain, however, how  $u(\cdot)$  can be derived solely from choices.

(Krantz et al. 1971). Because we can replace  $\Psi$  by any function which is a positive multiple of  $\Psi$ , we can replace  $Q(\cdot)$  by  $Q'(\cdot) = \alpha Q(\cdot)$ ,  $\alpha > 0$ . Because for all  $x$ ,  $\Psi(x,x) = 0$ ,  $Q(0) = 0$ .

Some authors use a slightly different expression for regret theory in which the utility of obtaining  $x$  and not  $y$  is equal to  $u(x) + R(u(x) - u(y))$ , where  $R$  is a regret-rejoice function, which is at least three times differentiable, strictly increasing, and decreasingly concave. This formulation of regret theory is equivalent to Eq. 2 when we define  $Q(u(x) - u(y)) = u(x) - u(y) + R(u(x) - u(y)) - R(u(y) - u(x))$ . Convexity of  $Q$  corresponds to  $R$  being decreasingly concave.

### 3. Elicitation Method

This Section explains our method for measuring regret theory. Even though our measurement method can be used to elicit the general regret model, Eq. 1, as we explain below, the intuition behind the method is clearer if we consider Eq. 2 and, hence, we will focus on this model.

Our method consists of two stages. In the first stage we use the tradeoff method (Wakker and Deneffe 1996) to elicit a *standard sequence* of outcomes  $\{x_0, \dots, x_k\}$  that are equally spaced in terms of utility, i.e. the elements of the standard sequence are such that  $u(x_{j+1}) - u(x_j) = u(x_1) - u(x_0)$  for all  $j$  in  $\{1, \dots, k-1\}$ . Hence, the first stage elicits the function  $u(\cdot)$ , which has the properties of Bell's (1982) incremental value function but without assuming the primitive of incremental value. In the second stage we use the standard sequence of outcomes to elicit the function  $Q(\cdot)$ .

It is easily verified, by substituting  $\Psi(x_{j+1}, x_j)$  for  $Q(u(x_{j+1}) - u(x_j))$  in the exposition below, that in the general regret theory of Loomes and Sugden (1987a) the standard sequence  $\{x_1, \dots, x_k\}$  is such that  $\Psi(x_{j+1}, x_j) = \Psi(x_1, x_0)$  for all  $j$  in  $\{1, \dots, k-1\}$ . The second stage then uses the standard sequence to elicit  $\Psi(\cdot, \cdot)$  for pairs of elements of the standard sequence. This shows that our method can indeed be used to elicit the general regret model, Eq. 1.

### 3.1. First stage: elicitation of $u(\cdot)$

We start by selecting the probability  $p$  with  $0 < p < 1$ , two *gauge outcomes*  $M$  and  $m$  with  $M > m$ , and a *starting outcome*  $x_0$ . Then we elicit the outcome  $x_1$  for which the decision maker is indifferent between  $(x_0, p; M)$  and  $(x_1, p; m)$ . The indifference  $(x_0, p; M) \sim (x_1, p; m)$  yields by Eq. 2:

$$pQ(u(x_0) - u(x_1)) + (1 - p) Q(u(M) - u(m)) = 0. \quad (3)$$

Rearranging and using the symmetry of  $Q(\cdot)$  gives:

$$Q(u(x_1) - u(x_0)) = \frac{1 - p}{p} Q(u(M) - u(m)). \quad (4)$$

We then determine the outcome  $x_2$  for which the decision maker is indifferent between  $(x_1, p; M)$  and  $(x_2, p; m)$ . Writing out this indifference gives by a similar line of argument as in Eqs. 3 and 4:

$$Q(u(x_2) - u(x_1)) = \frac{1 - p}{p} Q(u(M) - u(m)), \quad (5)$$

and, thus,  $Q(u(x_1) - u(x_0)) = Q(u(x_2) - u(x_1))$ . Because  $Q(\cdot)$  is strictly increasing, it follows that  $u(x_2) - u(x_1) = u(x_1) - u(x_0)$ . We proceed by eliciting indifferences  $(x_j, p; M) \sim (x_{j+1}, p; m)$  and in so doing we elicit a *standard sequence*  $\{x_1, \dots, x_k\}$  for which  $u(x_{j+1}) - u(x_j) = u(x_1) - u(x_0)$ ,  $j = 1, \dots, k-1$ . Because  $u(\cdot)$  is unique up to unit and location, we can arbitrarily choose the utility of two outcomes. We set  $u(x_0) = 0$  and  $u(x_k) = 1$ . It then follows that  $u(x_j) = j/k$  for  $j = 0, \dots, k$ .

### 3.2. Second stage: elicitation of $Q(\cdot)$

We use the standard sequence elicited in the first stage to measure  $Q(\cdot)$ .  $Q(\cdot)$  is scaled such that  $Q(u(x_1) - u(x_0)) = Q(1/k) = 1$ . Because  $u(x_{j+1}) - u(x_j) = 1/k$  for  $j = 0, \dots, k-1$ , it follows that  $Q(u(x_{j+1}) - u(x_j)) = 1$  for any two successive elements of the standard sequence. We considered two ways of measuring  $Q(\cdot)$ : a direct procedure and a chained procedure. In the *direct procedure*  $Q(j/k)$ ,  $j = 2, \dots, k$ , is measured by selecting an element  $x_j$  of the standard sequence and by eliciting the probability  $p$  for which the decision maker is indifferent between  $(x_0, p; x_1)$  and  $(x_j, p; x_0)$ . By (2) this gives:

$$\begin{aligned} pQ(u(x_0) - u(x_j)) + (1-p) Q(u(x_1) - u(x_0)) &= 0 \\ \Leftrightarrow pQ(j/k) &= 1 - p \end{aligned}$$

$$\Leftrightarrow Q(j/k) = \frac{1-p}{p}. \quad (6)$$

By making the standard sequence sufficiently fine, which can be achieved by appropriate selection of the two gauge outcomes, we can approximate  $Q(\cdot)$  to any desired degree of accuracy.

A disadvantage of the direct procedure is that the elicitation of  $p$  may be highly sensitive to response error when  $p$  is close to zero. A first problem is that even small response errors will then have strong effects on the elicited value of  $Q(\cdot)$ . This is most likely for values of  $Q(\cdot)$  close to  $Q(1)$ . For example, in our experiment we used  $k = 5$  and to measure  $Q(1)$  we must, by the above procedure, elicit the probability  $p$  for which a decision maker is indifferent between  $(x_0, p; x_1)$  and  $(x_5, p; x_0)$ . For our median subject this would mean that we had to elicit  $p$  such that indifference held between  $(\text{€}200, p; \text{€}390)$  and  $(\text{€}1220, p; \text{€}200)$ . Pretesting of our experiment showed that for most subjects this indifference only held for small probabilities  $p$ , leading to unstable estimates of  $Q(\cdot)$ .

A second problem associated with small indifference probabilities is that response errors can lead to biases in the elicited values of  $Q$ . Suppose that a subject's true indifference probability  $p$  is equal to 0.03. Then the room for error "to the left" of 0.03 is limited because for  $p = 0$ , the subject will, by dominance, have a clear preference for  $(x_0, p; x_1)$  over  $(x_j, p; x_0)$ . To the right of 0.03, there is, however, much more room for error and if the subject is equally likely to report a value lower than 0.03 as to report a value exceeding 0.03 then there will be an upward bias in the elicited indifference probability and, hence, a downward bias in  $Q(\cdot)$ .

Because of the above problems we decided to use a *chained procedure* for the determination of  $Q(\cdot)$ . The chained procedure measures  $Q(\cdot)$  by eliciting the probabilities  $p_j$  for which subjects are indifferent between  $(x_0, p_j; x_j)$  and  $(x_{j+1}, p_j; x_0)$ ,  $j = 1, \dots, k-1$ .<sup>3</sup> By Eq. 2 and the scaling of  $u(\cdot)$ , these indifferences yield:

$$p_j Q\left(\frac{j+1}{k}\right) = (1-p_j) Q\left(\frac{j}{k}\right), j = 1, \dots, k-1. \quad (7)$$

We know that  $Q(1/k) = 1$ , by the scaling of  $Q(\cdot)$ , and, hence, Eq. 7 measures  $Q(2/k)$  directly. Once we know  $Q(2/k)$ , we can use the value of  $Q(2/k)$  to determine  $Q(3/k)$ , etc. The chained procedure is less

---

<sup>3</sup> We could have used any other element of the standard sequence instead of  $x_0$  in the elicitation.

likely than the direct procedure to elicit probabilities close to zero, because it uses successive elements of the standard sequence. Pretesting showed that in the chained procedure most elicited probabilities were far away from 0, generally around 0.40.

A disadvantage of the chained method is that it may be sensitive to error propagation: errors made in the elicitation of  $Q(2/k)$  translate into errors in the elicitation of  $Q(3/k)$  and so on. To obtain insight into the extent to which our conclusions were affected by error propagation we performed simulation exercises, which are explained in Section 4 and the results of which are described in Section 5. The simulation exercises indicated that error propagation was not a major cause of concern in our study.

#### **4. Experiment**

The aim of the experiment was to elicit the functions  $u(\cdot)$  and  $Q(\cdot)$  through the procedure outlined in Section 3. This quantitative measurement also allowed testing for regret aversion, i.e. convex  $Q(\cdot)$ .

##### *Subjects and stimuli*

Subjects were 82 students (38 female, mean age 22) from different faculties of the Erasmus University Rotterdam. Before the actual experiment, we tested the design in 12 pilot sessions using other students and staff from INSEAD as participants. Subjects were paid a flat fee of €10 for their participation. In addition, one out of every ten subjects was selected randomly to play one of his choices for real at the end of the experiment. Regret theory predicts that this *random problem selection procedure* reveals subjects' true preferences in the sense that no differences will be observed with choices that are always played out for real. Empirical evidence provides support for this claim (Starmer and Sugden 1991, Beattie and Loomes 1997). A potential problem of using real incentives is that our method is based on chained measurements. This may induce subjects to respond strategically. For example, by giving too high values of  $x_i$  in the first stages of the elicitation of  $u(\cdot)$ , subjects could increase the attractiveness of the options that they faced in later stages of the elicitation. This danger is particularly present in matching tasks, in which subjects directly state their indifference value and this indifference value features in subsequent elicitation questions. We reduced the danger of strategic responses by using a choice-based

elicitation procedure. This makes it hard to notice the chained nature of the experiment because the derived indifference values do not appear in subsequent questions. To make the chained nature of the measurements even harder to discern, we included several filler questions that interspersed the elicitations.

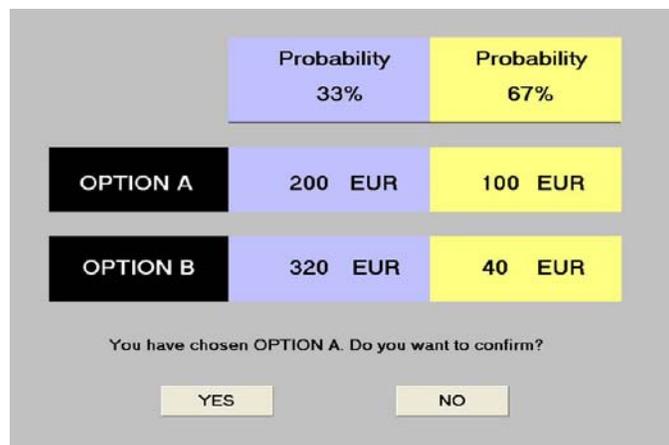
### *Procedure*

The experiment was run on a computer and was administered in small groups of at most 4 subjects. Subjects were seated in separate cubicles and no interaction between the subjects was possible. Two interviewers were present during each experimental session. Sessions lasted thirty minutes on average.

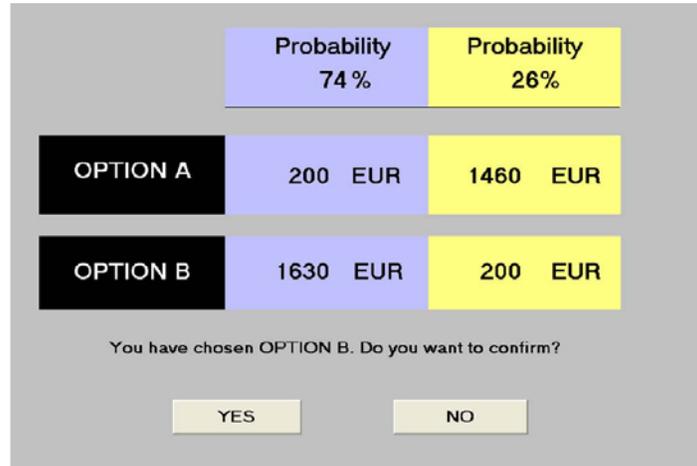
Each experimental session started with an explanation of the task, which was read aloud by one of the interviewers. Subjects could also see this explanation on their computer screen. Then a practice question followed. After this practice question the interviewer explained the random problem selection procedure. The introduction to the experiment ended with another practice question.

Figures 1 and 2 give examples of the questions in both stages of the elicitation. Subjects were presented with two options, labeled A and B, and they were asked to choose between the options by clicking on their preferred option, indifference was not allowed. They were then asked to confirm their choice. If they confirmed their choice they moved on to the next question. If not, they faced the same question again. The confirmation question intended to reduce the impact of response errors.

**Figure 1: Example of a screen faced in the first part of the experiment**



**Figure 2: Example of a screen faced in the second part of the experiment**



As mentioned before, we used a choice-based method to elicit indifference values. Empirical evidence suggests that choice-based procedures are less prone to inconsistencies than procedures that ask directly for indifference values (Bostic et al. 1990). When an indifference value had been determined through a series of choices, we repeated the first choice in the iteration process. If a subject made the same choice as before in this repeated question then the program moved on to the next choice-based elicitation. If not, the iteration process was started anew. Both options always yielded different outcomes under the two states of nature. This was deliberately done to avoid that event-splitting effects would confound the results.

To measure the function  $u(\cdot)$ , we elicited a standard sequence  $\{x_0, \dots, x_5\}$ . Hence,  $k = 5$  in our study. We set the probability  $p$  equal to  $\frac{1}{3}$ , the gauge outcomes  $M$  and  $m$  to €100 and €40, and the starting outcome  $x_0$  to €200. The probability  $\frac{1}{3}$  was chosen because the pilot sessions showed that subjects found it easy to understand. In the pilot sessions we set  $m$  equal to €0, but this produced strong risk aversion: a substantial proportion of subjects chose  $(€200, \frac{1}{3}; €100)$  over  $(€x_1, \frac{1}{3}; €0)$  even for high values of  $x_1$ .

As we explained in Section 3, our measurement method amounts to finding values  $x_{j+1}$  for which  $(€x_j, \frac{1}{3}; €100) \sim (€x_{j+1}, \frac{1}{3}; €40)$ ,  $j = 0, \dots, 4$ . In the first iteration of the choice-based elicitation procedure, we chose  $x_{j+1}$  so that the two acts  $(€x_j, \frac{1}{3}; €100)$  and  $(€x_{j+1}, \frac{1}{3}; €40)$  had equal expected value. Depending on the subject's choice,  $x_{j+1}$  was increased or decreased. After a switch in preference, the change in  $x_{j+1}$  was reduced and the direction of change was reversed. In this way, we zoomed in on the subject's

indifference value of  $x_{j+1}$ . The exact procedure followed in this “zooming-in-process” is explained in Appendix A. The number of iterations per elicitation depended on the subject’s choices. The probabilities  $p$  were kept fixed at  $\frac{1}{3}$  in the elicitation of the standard sequence. To stimulate subjects paying attention to the probability dimension, we varied the probabilities of the two states in the filler questions that interspersed the elicitation questions.

Two problems could arise in the elicitation of  $x_{j+1}$ . In case a subject violated *monotonicity*, i.e. one of the options gave at least as good an outcome under both states of nature, he was given a warning that explained that monotonicity was violated and he was asked to reconsider his choice. If the subject persisted in his choice, the experiment ended. Another problem that could arise was that a subject always chose  $(\text{€}x_j, \frac{1}{3}; \text{€}100)$  over  $(\text{€}x_{j+1}, \frac{1}{3}; \text{€}40)$ , even when  $x_{j+1}$  was much larger than  $x_j$ , and indifference between the two options would not be reached. To avoid this problem, the experiment ended when a subject still preferred  $(\text{€}x_j, \frac{1}{3}; \text{€}100)$  to  $(\text{€}x_j + \text{€}820, \frac{1}{3}; \text{€}40)$ . In the pilots we observed such extreme preferences once.

The second stage determined four values of  $Q(\cdot)$ :  $Q(2/5)$ ,  $Q(3/5)$ ,  $Q(4/5)$ , and  $Q(1)$  through the chained procedure. We measured  $Q(\cdot)$  by eliciting the probabilities  $p_{j+1}$  for which subjects were indifferent between  $(\text{€}200, p_{j+1}; \text{€}x_j)$  and  $(\text{€}x_{j+1}, p_{j+1}; \text{€}200)$ ,  $j = 1, \dots, 4$ . The elicitation procedure was similar to that used in the elicitation of  $u(\cdot)$  and is explained in Appendix A. In contrast with the elicitation of  $u(\cdot)$ , the choice-based elicitations did not have to be performed consecutively in the elicitation of  $Q(\cdot)$ . Hence, we could intersperse the different iteration processes, thereby obscuring that we were looking for indifference values.

It was possible that subjects violated stochastic dominance in the second part of the experiment. For any acts  $f$  and  $g$ ,  $f$  *stochastically dominates*  $g$  if for all outcomes  $x$ , the probability that  $f$  generates an outcome at least as preferred as  $x$  is at least as great as the corresponding probability for  $g$  and for some outcome  $x$  the probability for  $f$  exceeds the probability for  $g$ . For example,  $f = (\text{€}400, \frac{2}{3}; \text{€}200)$  stochastically dominates  $g = (\text{€}200, \frac{2}{3}; \text{€}300)$ . The program kept track of the number of violations of stochastic dominance.

Three consistency checks completed the experiment. We repeated the elicitation of  $p_4$  to examine consistency in subjects' responses. We also elicited the probabilities  $p_{c1}$  and  $p_{c2}$  that made subjects indifferent between  $(\text{€}x_3, p_{c1}; \text{€}x_2)$  and  $(\text{€}x_4, p_{c1}; \text{€}200)$  and between  $(\text{€}x_1, p_{c2}; \text{€}x_3)$  and  $(\text{€}x_2, p_{c2}; \text{€}200)$ .<sup>4</sup> The indifference between  $(\text{€}x_3, p_{c1}; \text{€}x_2)$  and  $(\text{€}x_4, p_{c1}; \text{€}200)$  yields after some rearranging and using the symmetry and scaling of  $Q(\cdot)$ :

$$Q(2/5) = \frac{p_{c1}}{1 - p_{c1}}. \quad (8)$$

For consistency we should observe that the value of  $Q(2/5)$  in Eq. 8 was equal to the value of  $Q(2/5)$  elicited before. That is, we should observe that  $\frac{p_{c1}}{1 - p_{c1}} = \frac{1 - p_2}{p_2}$ , i.e. that  $p_{c1} = 1 - p_2$ . The indifference between  $(\text{€}x_1, p_{c2}; \text{€}x_3)$  and  $(\text{€}x_2, p_{c2}; \text{€}200)$  yields:

$$Q(3/5) = \frac{p_{c2}}{1 - p_{c2}}. \quad (9)$$

For consistency we should observe that the value of  $Q(3/5)$  in Eq. 9 was equal to the value of  $Q(3/5)$  elicited before, i.e.  $\frac{p_{c2}}{1 - p_{c2}} = \frac{1 - p_3}{p_3} \frac{1 - p_2}{p_2}$  or  $p_{c2} = \frac{(1 - p_3)(1 - p_2)}{1 - p_3 - p_2 + 2p_3p_2}$ . Systematic inequalities would indicate violations of regret theory.

### *Analyses*

We present individual and aggregate data for  $u(\cdot)$  and  $Q(\cdot)$ . In the analysis of the aggregate data, we focus on the results based on the medians and only report the results based on the means in case they led to different conclusions. Throughout the analyses, significance of differences was tested both parametrically (paired t-test) and nonparametrically (Wilcoxon test). Because these tests always led to the same conclusion, we only report the nonparametric results. Differences between proportions were tested by the binomial test.

---

<sup>4</sup> The c in  $p_{c1}$  and  $p_{c2}$  serves as a reminder that these are the responses elicited in the consistency questions.

### Analysis of $u(\cdot)$

To investigate curvature of  $u(\cdot)$  at the individual level, we computed for each subject the differences

$$\Delta_{gh,\ell m} = (x_g - x_h) - (x_\ell - x_m), g > h, g > \ell, g - h = \ell - m, \quad (10)$$

with all  $x_j, j = g, h, \ell, m$ , elements of the elicited standard sequence. Because  $g - h = \ell - m$ , it follows from the construction of the standard sequence that  $u(x_g) - u(x_h) = u(x_\ell) - u(x_m)$ . Hence, a positive value of  $\Delta_{gh,\ell m}$  corresponds to a concave part of  $u(\cdot)$ , a negative value to a convex part, and a value of zero to a linear part. We could observe 20 values of  $\Delta_{gh,\ell m}$ . To account for response error, we classified a subject's utility function as concave/convex/linear if at least 12 out of 20 values of  $\Delta_{gh,\ell m}$  were positive/negative/zero.<sup>5</sup> Otherwise, a subject was classified as mixed.

To smoothen out irregularities in the data, we also analyzed the data under specific parametric assumptions about utility. We examined two parametric families of utility, the power family and the exponential family. Both families are widely used in economics and decision theory. Let  $a = 1/(x_5^r - 200^r)$  and  $b = 200^r/(x_5^r - 200^r)$ . The *power family* was defined by  $ax^r - b$  if  $r > 0$ , by  $\ln(x) - b$  if  $r = 0$ , and by  $-(ax^r + b)$  if  $r < 0$ . This scaling entailed that  $u(200) = 0$  and  $u(x_5) = 1$ . It is well known that  $r < 1$  corresponds to concave utility,  $r > 1$  to convex utility, and  $r = 1$  to linear utility. Let  $z = (x - x_0)/(x_5 - x_0)$ ,  $x \in [200, x_5]$ . The *exponential family* was defined by  $(e^{rz} - 1)/(e^r - 1)$  if  $r \neq 0$  and by  $z$  if  $r = 0$ . The case  $r < 0$  corresponds to concave utility and the case  $r > 0$  to convex utility. We estimated the power family and the exponential family by nonlinear least squares both for the median data and for each subject separately.

We used the parametric estimates to obtain two other, parametric, classifications of the subjects. First, we classified a subject as concave if his power coefficient was below 0.95, as linear if his power coefficient was between 0.95 and 1.05, and as convex if his power coefficient exceeded 1.05. Second, using the standard errors of the coefficients, we classified a subject as concave if his power coefficient was statistically significantly smaller than 1 and as convex if his power coefficient was statistically significantly greater than 1. Here we did not use the linear classification because a coefficient that does not

---

<sup>5</sup> Similar criteria were used by Fennema and van Assen 1999, Abdellaoui 2000, and Etchart-Vincent 2004.

significantly differ from 1 does not imply that a linear function fits the data particularly well. For the exponential function we used the classifications “concave if the exponent was less than  $-0.05$ ”, “linear if the exponent was between  $-0.05$  and  $0.05$ ”, and “convex if the exponent was greater than  $0.05$ ” and “concave if the exponent was significantly less than 0” and “convex if the exponent was significantly greater than 0”. The results were very similar for the power and the exponential family. To facilitate comparability of our result with earlier studies on utility measurement, we report the results for the power family in what follows. The results for the exponential family can be found in the electronic companion to this paper.

#### *Analysis of $Q(\cdot)$*

To investigate curvature of  $Q(\cdot)$  at the individual level, we computed the differences

$$\nabla_{gh,\ell m} = (Q(g/5) - Q(h/5)) - (Q(\ell/5) - Q(m/5)), \quad g > h, \quad g > \ell, \quad g - h = \ell - m, \quad (11)$$

with  $g, h, \ell, m \in \{0, \dots, 5\}$ . Because  $g - h = \ell - m$ , it follows that a positive value of  $\nabla_{gh,\ell m}$  corresponds to a convex part of  $Q(\cdot)$ , a negative value to a concave part, and a value of zero to a linear part. We could observe 20 values of  $\nabla_{gh,\ell m}$ . We classified a subject’s  $Q(\cdot)$  as concave if at least 12 out of 20 values of  $\nabla_{gh,\ell m}$  were positive and as convex if at least 12 out of 20 values were negative. For reasons that will become apparent below, we also studied the behavior of  $Q(\cdot)$  on  $[1/5, 1]$ . To do so, we considered the subset of values of  $\nabla_{gh,\ell m}$  for which  $m > 0$ . Then we could observe 10 values of  $\nabla_{gh,\ell m}$  and we used a cut-off value of 6 out of 10.

We also fitted the power family and the exponential family to  $Q(y)$ ,  $y \in [0, 1]$  and to the subdomains  $y \in [0, 2/5]$  and  $y \in [1/5, 1]$ . For both families  $Q(\cdot)$  was rescaled to ensure that  $Q(1) = 1$  in the estimations on  $[0, 1]$  and on  $[1/5, 1]$ . In the estimation of  $Q(y)$  on  $[1/5, 1]$  we transformed the interval  $[1/5, 1]$  to the unit interval by defining  $z = (y - 1/5)/(4/5)$  and  $Q'(z) = (Q(y) - 1)/(Q(1) - 1)$ . To determine the power and exponential coefficients on  $[0, 2/5]$ , we rescaled  $Q$  such that  $Q(2/5) = 1$ . The power and exponential coefficients were uniquely determined on  $[0, 2/5]$  because there was only one degree of freedom for  $Q$ , the value of  $Q(1/5)$ , on this subdomain.

As for utility, we used the parametric fittings to obtain two other classifications of subjects based on the shape of  $Q(\cdot)$ , the  $\pm 0.05$  classification and the classification based on statistical significance. On  $[0, \frac{2}{5}]$  the power and exponential coefficients were uniquely determined and, we could not use the classification of subjects based on statistical significance.

For  $Q$  we focus on the exponential estimates, because these led to better convergence. As explained by Wakker (2006), problems may arise with the power function when zero is included in the domain. The exponential function does not have these problems around zero. The results for the power function, which were close to those obtained for the exponential function, can be found in the electronic companion.

#### *Analysis of error propagation*

In Section 3 we discussed the danger of error propagation affecting the results. To test for the impact of error propagation both in the elicitation of  $u(\cdot)$ , where previously elicited elements of the standard sequence were used in the elicitation of successive elements, and in the elicitation of  $Q(\cdot)$  by the chained method, we adopted the error model that was proposed by Fechner (1860/1966). The Fechner model was used by Hey and Orme (1994) for decision theory and has also been successfully applied in behavioral game theory (McKelvey and Palfrey 1995, Goeree et al. 2003). Essentially, the idea is that subjects make errors in computing the true difference in utility between options and that actual choices on any particular occasion are made according to the perceived difference in utility, which is equal to the true utility difference plus some noise. Two alternative error models have been proposed. The trembling hand model (Harless and Camerer 1994) argues that subjects make mistakes in reporting their choices. The random preference model (Loomes and Sugden 1995) argues that people are uncertain about their utility function and pick one such function at random from a set of utility functions. The trembling hand model was found to be inconsistent with empirical evidence (Loomes and Sugden 1998, Loomes 2005), while the random preference model was hard to apply to our data. Hence, we did not consider these two error models.

We operationalized the Fechner model as follows. In the elicitation of  $u(\cdot)$  we assumed that the subject could make an error in the computation of the utility difference between  $x_{j+1}$  and  $x_j$ . Recall that we scaled utility such that  $u(x_{j+1}) - u(x_j) = 0.20$ . The Fechner model implies that in the each elicitation the actual utility difference between  $x_{j+1}$  and  $x_j$  was equal to the assumed value 0.20 and some error  $\varepsilon_{j+1}$ ,  $j = 0, \dots, 4$ :

$$u(x_{j+1}) - u(x_j) = 0.20 + \varepsilon_{j+1}. \quad (12)$$

Because  $u(x_0) = 0$ , we obtain  $u(x_1) = 0.20 + \varepsilon_1$ . From Eq. 12,  $u(x_2) = u(x_1) + 0.20 + \varepsilon_2 = 0.40 + \varepsilon_1 + \varepsilon_2$ . In general, our error specification gives

$$u(x_{j+1}) = \frac{j+1}{5} + \sum_{i=1}^{j+1} \varepsilon_i, \quad j = 0, \dots, 4. \quad (13)$$

In the elicitation of  $Q(\cdot)$  we assumed that

$$\begin{aligned} Q(u(x_{j+1})) &= Q\left(\frac{j+1}{5} + \sum_{i=1}^{j+1} \varepsilon_i\right) = \frac{1-p_{j+1}}{p_{j+1}} Q(u(x_j)) + \phi_{j+1} = \\ &= \frac{1-p_{j+1}}{p_{j+1}} Q\left(\frac{j}{5} + \sum_{i=1}^j \varepsilon_i\right) + \phi_{j+1}, \quad j = 1, \dots, 4, \end{aligned} \quad (14)$$

where we denote the error in the elicitation of  $Q(\cdot)$  by  $\phi$  to distinguish it from the error in the elicitation of  $u(\cdot)$ . Note that the error in the elicitation of  $u(\cdot)$  affects the elicitation of  $Q(\cdot)$ . Because  $Q(u(x_1)) = 1$ , we get

$$Q(u(x_2)) = \frac{1-p_2}{p_2} + \phi_2. \text{ Substituting this expression in the expression for } Q(u(x_3)) = \frac{1-p_3}{p_3} Q(u(x_2)) + \phi_3$$

gives  $Q(u(x_3)) = \frac{1-p_3}{p_3} \left(\frac{1-p_2}{p_2} + \phi_2\right) + \phi_3$ . In general we obtain:

$$Q(u(x_{j+1})) = \prod_{i=2}^{j+1} \frac{1-p_i}{p_i} + \sum_{i=2}^{j+1} \left(\phi_i \prod_{k=i+1}^{j+1} \frac{1-p_k}{p_k}\right) + \phi_{j+1}. \quad (15)$$

We examined two different error specifications. In the first specification, we assumed that the errors  $\varepsilon_j$  and  $\phi_j$  were normally distributed with mean 0 and constant standard deviation of 0.05. The size of the standard deviation is not important as the purpose of our analysis was to examine whether small response errors were propagated into large errors in the elicited values of  $u(\cdot)$  and  $Q(\cdot)$ . The assumption of constant variance may be too restrictive as it is likely that subjects are more error-prone the larger the difference between the outcomes. We therefore also examined a model in which the errors were

heteroskedastic and the size of the errors varied with the difference between the outcomes of the standard sequence. In this model the standard deviation of the errors was assumed equal to  $\sigma(x_{j+1} - x_j)$  where we normalized such that the standard deviation for the median subject was 0.05 in the first elicitation of the standard sequence. In contrast with the constant error specification, the heteroskedastic error specification excludes violations of monotonicity, which were not allowed in our final data. We performed 1000 simulations both for  $u(\cdot)$  and for  $Q(\cdot)$  and for both error specifications.

## 5. Results

Sixteen subjects were excluded because they were extremely risk averse (6 in the elicitation of  $x_1$ , 5 in the elicitation of  $x_3$ , 2 in the elicitation of  $x_4$ , and 3 in the elicitation of  $x_5$ ). They chose  $(\epsilon x_j, \frac{1}{3}; \epsilon 100)$  over  $(\epsilon x_j + 820, \frac{1}{3}; \epsilon 40)$  for some  $j = 0, \dots, k-1$ . Two subjects were excluded because they violated monotonicity even after this violation had been pointed out to them by the program, and eight subjects were excluded because they violated stochastic dominance. This left 56 subjects (22 female, mean age 22.1) in the analysis. Note that the 2.4% of subjects violating monotonicity is in line with the results of Loomes and Sugden (1998). Moreover, having 10.7% of the subjects violate stochastic dominance when stochastic dominance is not immediately obvious is not disturbingly high compared to other studies. An overview of the elicited results for the subjects included in the analyses is in Appendix B.

The subjects who were excluded because they chose  $(\epsilon x_j, \frac{1}{3}; \epsilon 100)$  over  $(\epsilon x_j + 820, \frac{1}{3}; \epsilon 40)$  for some  $j = 0, \dots, k-1$ , in the elicitation of  $u(\cdot)$  were those who were most risk averse. Excluding these subjects might bias the results and, therefore, we also analyzed the results while excluding the sixteen most risk seeking subjects. To determine the risk attitude of a subject, we recorded how often the value of  $x_{j+1}$  that yielded indifference between  $(\epsilon x_j, \frac{1}{3}; \epsilon 100)$  and  $(\epsilon x_{j+1}, \frac{1}{3}; \epsilon 40)$  was less than  $x_j + 120$ , the value that ensured that the two acts had equal expected value. Such an indifference value is consistent with risk-seeking behavior (Rothschild and Stiglitz 1970). It turned out that excluding the most risk seeking subjects did not affect the main conclusions and we will, therefore, not discuss these results in what follows. They can be found in the electronic companion.

### *Consistency tests*

The mean and median values of  $p_4$  in the actual experiment were nearly equal to those in the retest (the medians were both 0.42, the means were 0.415 and 0.417) and there was no significant difference between the two elicitations of  $p_4$  ( $p = 0.767$ ). We also observed no significant difference in the second consistency test, i.e. between  $p_{c2}$  and  $\frac{(1 - p_3)(1 - p_2)}{1 - p_3 - p_2 + 2p_3p_2}$ , ( $p = 0.205$ ). In the first consistency test, we observed, however, a significant difference between the median values of  $p_{c1}$  and  $(1 - p_2)$  (0.740 versus 0.595,  $p < 0.01$ ). We discuss this difference in further detail in Section 7.

Recall from Section 4, that after each choice-based iteration process we repeated the first choice of this iteration process and that the iteration process was started anew if the subject did not make the same choice. For  $u(\cdot)$ , 78.2% of the elicitations required only one iteration process, 12.9% of the elicitations required the iteration process to be repeated once, and 8.9% of the elicitations required the iteration process to be repeated twice. In the elicitation of  $Q(\cdot)$ , these proportions were 70.5, 16.1, and 13.4. The data do not suggest important differences in volatility across the elicitations.

### *The elicitation of $u(\cdot)$*

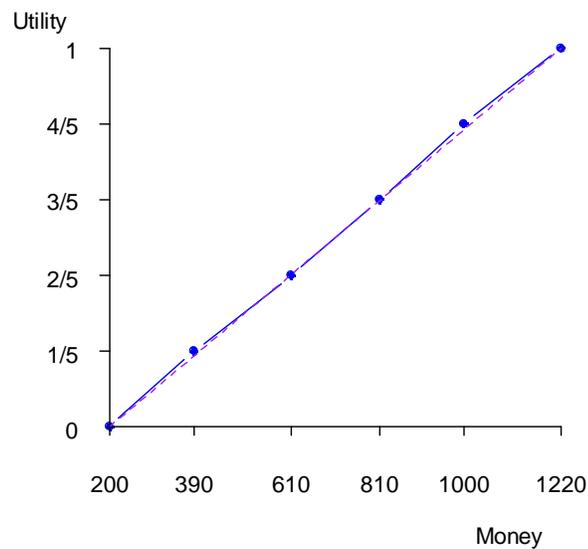
Concavity of  $u(\cdot)$  was the dominant pattern at the individual level with 29 subjects having concave utility, 11 having convex utility, and 2 having linear utility. The remaining 14 subjects could not be classified. The proportion of concave subjects was significantly higher than the proportion of convex subjects ( $p < 0.01$ ).

The classification of the subjects based on their fitted power coefficients confirmed that concave utility was more common than convex utility. In the first parametric classification of subjects, the  $1 \pm 0.05$  classification, 29 had concave utility, 19 had convex utility, and 8 had linear utility; in the second, the classification based on statistical significance, 22 had concave utility and 13 had convex utility. In both parametric classifications, the proportion of concave subjects was only marginally higher than the proportion of convex subjects ( $p = 0.083$  in the  $1 \pm 0.05$  classification and  $p = 0.064$  in the classification

based on statistical significance). The median of the individual estimates of the power coefficient was equal to 0.94.

Our method for eliciting utility is also valid under prospect theory (Kahneman and Tversky 1979, Tversky and Kahneman 1992). The proportions of concave, linear, and convex subjects that we observed were comparable to those observed in other studies using the tradeoff method (Abdellaoui 2000, Abdellaoui et al. 2005).

**Figure 3: The elicited utility function based on the median data**



We also analyzed the data at the aggregate level. In many decision contexts policy makers are interested in aggregate data. An important example is the economic evaluation of health care, where recommendations have to be made about the reimbursement of new treatments, and where typically the preferences of a representative sample from the general population are elicited and the interest is in the mean or median preferences. At the aggregate level utility was close to linear. Figure 3 shows the shape of the utility function  $u(\cdot)$  for the median data. The dotted line is drawn for comparison and represents the case of linear utility. Consistent with linearity of  $u(\cdot)$ , we observed no significant differences between  $(x_{j+1}-x_j)$  and  $(x_j-x_{j-1})$  for  $j = 1, \dots, 4$  ( $p > 0.10$  in all pairwise tests). Parametric fitting also confirmed that

$u(\cdot)$  was close to linear: the power coefficient based on the median data was equal to 0.96 and was not significantly different from 1, the value corresponding to linear utility ( $p = 0.106$ ). The power coefficient based on the mean data was equal to 0.88, indicating more concavity, and was significantly different from 1 ( $p < 0.01$ ).

Compared with earlier papers that measured utility under prospect theory, our aggregate findings were close to those reported in Tversky and Kahneman (1992), Abdellaoui (2000), Abdellaoui et al. (2005), Schunk and Betsch (2006), and one of the estimates in Abdellaoui et al. (forthcoming-a) where the estimates for the power coefficient varied between 0.88 and 0.91. The estimates of Gonzalez and Wu (1999), Abdellaoui et al. (forthcoming-b), and two other estimates in Abdellaoui et al. (forthcoming-a) were, however, lower; they varied between 0.49 and 0.77 and indicated more concavity of utility.

*The elicitation of  $Q(\cdot)$*

Table 1 displays the elicited median probabilities and their interquartile ranges. The table shows that the probabilities were generally far from zero and, hence, the problems associated with small probabilities that we discussed in Section 3 did not cause serious problems. There were only three probabilities less than 0.15 in the data

**Table 1: The elicited probabilities**

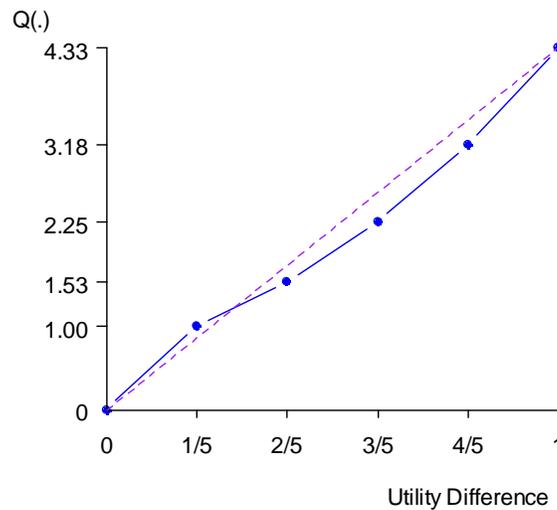
	Median	Interquartile range
$p_2$	0.40	0.32-0.44
$p_3$	0.40	0.38-0.43
$p_4$	0.42	0.40-0.45
$p_5$	0.43	0.40-0.46

At the individual level, convexity was the dominant pattern with 30 subjects being classified as convex and 12 as concave. The remaining 14 subjects displayed a mixed shape for the regret function. No subject was classified as linear, the case corresponding to expected utility. The proportion of convex subjects was significantly higher than the proportion of concave subjects ( $p < 0.01$ ).

Convexity of  $Q(\cdot)$  was also the dominant pattern when we looked at the individual parametric estimates. In the  $0 \pm 0.05$  classification, 33 subjects were classified as convex, 21 as concave, and 2 as linear. In the classification based on statistical significance, 22 subjects were convex and 13 were concave. In both classifications the difference between convex and concave was only marginally significant ( $p = 0.053$  and  $p = 0.062$ , respectively). The median of the individual exponential estimates was equal to 0.19.

The conclusion of predominant convexity needs to be qualified, however. Convexity requires that  $Q(2/5) > 2*Q(1/5)$  and, hence, that  $p_2 < 1/3$ . There were only 15 subjects for whom this was true suggesting that for most subjects  $Q(\cdot)$  started off concave. The proportion of subjects for whom  $p$  exceeded  $1/3$  was significantly larger than the proportion for whom  $p$  was less than  $1/3$  ( $p < 0.01$ ).

**Figure 4: The elicited function  $Q(\cdot)$  based on the median data**



The initial concavity of  $Q(\cdot)$  was confirmed when we looked at the aggregate data. Figure 4 displays the shape of  $Q(\cdot)$  for the median data. The dotted line serves as the benchmark of linear  $Q(\cdot)$ . The figure shows that  $Q(\cdot)$  was generally convex except for the first part of the function where it was concave. Consistent with concavity on  $[0, 2/5]$ ,  $Q(2/5) - Q(1/5)$  was significantly smaller than  $Q(1/5) - Q(0) = 1$  ( $p < 0.01$ ). Consistent with convexity of  $Q(\cdot)$  on  $[1/5, 1]$ ,  $Q((j+1)/5) - Q(j/5)$  was always significantly larger than

$Q(j/5) - Q((j-1)/5)$ ,  $j = 2, \dots, 4$ , ( $p < 0.01$  in all pairwise tests). The exponential coefficient based on the median data was 0.45, indicating slight convexity of  $Q(\cdot)$ . The exponent was significantly different from 0, the value corresponding to expected utility ( $p = 0.043$ ).<sup>6</sup>

#### *Results on subdomains*

Because the above analysis suggests that the curvature of  $Q$  is different on  $[0, \frac{2}{5}]$  than on  $[\frac{1}{5}, 1]$ , we analyzed these two subdomains separately. On  $[0, \frac{2}{5}]$ , 41 subjects were classified as concave and 15 as convex ( $p < 0.01$ ). On  $[\frac{1}{5}, 1]$ , 42 subjects were classified as convex and only 9 as concave with the remaining 5 subjects classified as mixed. The proportion of convex subjects was significantly higher than the proportion of concave subjects ( $p < 0.01$ ).

The parametric classifications were similar to the nonparametric classifications. On  $[\frac{1}{5}, 1]$ , 44 subjects were convex, 8 were concave and 4 were linear ( $p < 0.01$ ) based on the  $0 \pm 0.05$  classification. Based on the statistical significance classification 38 subjects were classified as convex and 6 as concave ( $p < 0.01$ ). The median of the individual exponential coefficients was equal to 1.13. On  $[0, \frac{2}{5}]$ , the median of the individual exponential coefficients was  $-1.28$ .<sup>7</sup>

On  $[0, \frac{2}{5}]$ , the exponential based on the median data was also  $-1.28$ . On  $[\frac{1}{5}, 1]$ , the exponential estimate based on the median data was 1.01, which was significantly different from 0 and indicated more pronounced convexity on  $[\frac{1}{5}, 1]$  than on  $[0, 1]$ .

#### *Results on error propagation*

Tables 2 and 3 show the effect of error propagation on the results. The Tables display the effects of the homoskedastic error specification on the elicited values of  $u(\cdot)$  and  $Q(\cdot)$ . The results under the heteroskedastic error specification were similar and can be found in the electronic companion.

The effect of error propagation was small. Remember that we ran 1000 simulations to analyze the impact of error propagation. For each simulation we computed the medians of the individual values of

---

<sup>6</sup> When the 16 most risk-seeking subjects were excluded the exponent based on the median data did not differ significantly from 0, however ( $p = 0.716$ ).

<sup>7</sup> The classification of subjects was the same as the nonparametric classification because we computed rather than estimated the power coefficient on  $[0, \frac{2}{5}]$ .

$u(x_j)$  and  $Q(j/5)$ ,  $j=1, \dots, 5$ . Tables 2 and 3 show the medians of these 1000 medians (second column), their standard deviations (third column), and their interquartile ranges (fourth column). The tables show that introducing errors does not lead to large variation in the elicited values of  $u(x_j)$  and  $Q(j/5)$ ,  $j=1, \dots, 5$ . Even though the standard deviations increased over the elicitation of  $u(\cdot)$  they remained modest compared to the initial standard deviation of the error term of 0.05. This was also reflected in the interquartile ranges, which were small and always less than 0.05. For  $Q(\cdot)$ , the standard deviations and the interquartile ranges were somewhat larger but they remained comparable to the initially selected standard deviation of 0.05, suggesting that error propagation was no major problem in the elicitation of  $Q(\cdot)$  either.

**Table 2: Impact of error propagation on  $u(\cdot)$**

	Median	Standard deviation	Interquartile range
$u(x_1)$	0.200	0.008	0.194-0.205
$u(x_2)$	0.400	0.012	0.392-0.409
$u(x_3)$	0.600	0.014	0.589-0.610
$u(x_4)$	0.800	0.016	0.788-0.810
$u(x_5)$	0.999	0.018	0.986-1.012

To illustrate the modest impact of error propagation on the measurement of  $Q(\cdot)$  we re-estimated the optimal exponential coefficients for  $Q(\cdot)$  by combining the 25% interquartile ranges of  $u(x_1)$ ,  $u(x_2)$ , and  $u(x_3)$  with the 75% interquartile ranges of  $Q(u(x_1))$ ,  $Q(u(x_2))$ , and  $Q(u(x_3))$  and the 75% interquartile ranges of  $u(x_4)$  and  $u(x_5)$  with the 25% interquartile ranges of  $Q(u(x_4))$  and  $Q(u(x_5))$ .<sup>8</sup> This “bad case scenario” will make the elicited shape of  $Q(\cdot)$  more concave. As a result, the estimated exponential coefficients fell slightly from 0.45 to 0.29 on  $[0,1]$  and from 1.01 to 0.89 on  $[\frac{1}{5},1]$ . On  $[0,1]$  the exponential coefficient was no longer significantly different from 0 ( $p = 0.170$ ).

<sup>8</sup> That is we based the estimations on (0.194,1), (0.392,1.561), (0.589,2.282), (0.810,3.118), and (1.012,4.287).

**Table 3: Impact of error propagation on  $Q(\cdot)$** 

	Median	Standard deviation	Interquartile range
$Q(2/5)$	1.547	0.022	1.532-1.561
$Q(3/5)$	2.257	0.036	2.236-2.282
$Q(4/5)$	3.159	0.060	3.118-3.197
$Q(1)$	4.343	0.085	4.287-4.398

## 6. Predictions and Implications

In this Section we will elaborate on some predictions and implications of our findings. First we will show that our findings predict a well-known violation of expected utility, the common consequence effect. Kahneman and Tversky (1979) observed the following preferences. In a comparison between prospect A giving 2400 Israeli pounds for sure and prospect B giving 2500 Israeli pounds with probability 0.33, 2400 Israeli pounds with probability 0.66 and nothing with probability 0.01, 82% of their subjects preferred A. In a comparison between prospect C giving 2400 Israeli pounds with probability 0.34 and nothing otherwise and prospect D giving 2500 Israeli pounds with probability 0.33 and nothing otherwise, 83% of their subjects preferred D. It is well-known that these preferences violate expected utility. When we substituted Euros for pounds<sup>9</sup>, our estimates implied that 54% of our subjects preferred A to B and 64% preferred D to C.<sup>10</sup> Hence, our estimates predict the common consequence effect. That we observe the common consequence effect to a lesser degree than Kahneman and Tversky should not come as a surprise. In the elicitation of  $u(\cdot)$  and  $Q(\cdot)$  we used only questions in which both prospects are risky. It is well-known that deviations from expected utility are less pronounced when both prospects are risky (Starmer 1992) than when one of the prospects is riskless as in the comparison between prospects A and B.

<sup>9</sup> This seems reasonable as 2500 Israeli pounds was close to a family's monthly income at the time of Kahneman and Tversky's study and so was 2500 Euros at the time of our study.

<sup>10</sup> For some subjects we had to extrapolate outside the domain within which we had elicited  $u(\cdot)$ . We selected these pairs of prospects because we wanted to check whether we could predict preferences that had been observed in the literature and for these pairs the need for extrapolation was modest compared with other pairs that have been used in the literature.

A distinguishing feature of regret theory is that it allows for intransitive preferences. We show next that our estimates are consistent with intransitive preference cycles. Consider our median data, which can be interpreted as representing the representative subject from this study, and the following three prospects with  $\varepsilon > 0$ :

Probability Act	10%	6%	10%	6%	68%
A	810	810	200	200	200
B	610	610	610	200	200
C	610	200	$610 - \varepsilon$	610	200

For our median data, we have  $A > B$  and, provided that  $\varepsilon$  is small,  $C > A$ .<sup>11</sup> Obviously,  $B > C$  and, hence, we obtain a preference cycle. This preference cycle is similar to the one observed by Starmer (1999) and cannot be explained by most of the other nonexpected utility models, including new prospect theory (Tversky and Kahneman (1992)). The only theory that can also explain this cycle is original prospect theory (Kahneman and Tversky (1979)).

In our study convexity of  $Q(\cdot)$  leads to less risk averse behavior. To see this, consider the elicitation of  $Q(\cdot)$ . There we determined the probability  $p_{j+1}$  for which subjects were indifferent between  $(\text{€}200, p_{j+1}, \text{€}x_j)$  and  $(\text{€}x_{j+1}, p_{j+1}, \text{€}200)$ . By stochastic dominance the indifference probability must be less than  $\frac{1}{2}$  and the closer it is to  $\frac{1}{2}$ , the more risk averse is a subject. Let  $p_{j+1}^*$  denote the probability corresponding to linear  $Q$ , i.e. expected utility. To have convex  $Q(\cdot)$  the indifference probability must be less than  $p_{j+1}^*$ . Consequently, convex  $Q(\cdot)$  entails a shift of the indifference probability away from  $\frac{1}{2}$  and, hence, less risk aversion. That the presence of regret leads to less risk averse behavior was also observed by Larrick and Boles (1995) in an experimental study on negotiation. Interestingly, Larrick and Boles (1995) concluded that the presence of regret might improve decision making in organizations because

<sup>11</sup> For our median data,  $u(810) - u(610) = 1/5$ ,  $u(610) - u(200) = 2/5$ ,  $Q(3/5) = 2.25$ ,  $Q(2/5) = 1.53$ , and  $Q(1/5) = 1$ . We obtain that  $0.16 * Q(u(810) - u(610)) - 0.10 * Q(u(610) - u(200)) = 0.07 > 0$ , and, thus,  $A > B$ . If  $\varepsilon = 0$ ,  $0.10 * Q(u(810) - u(610)) + 0.06 * Q(u(810) - u(200)) - 0.16 * Q(u(610) - u(200)) = -0.01$  and, therefore,  $C > A$ . If  $\varepsilon$  is sufficiently small then  $C > A$  will still hold.

organizational researchers typically find that organizations are too risk averse and recommend that they encourage risk neutral attitudes.

Let us finally turn to the relevance of our findings for field data. A well known phenomenon in finance is that investors appear reluctant to realize losses and eager to realize gains. This disposition effect, as it is called, has been observed for stocks (Grinblatt and Keloharju 2001), the exercise of options (Heath et al. 1999), and the sale of residential housing (Genesove and Mayer 2001) and cannot be explained by portfolio rebalancing and transaction costs (Odean 1998). Much of the literature has suggested that the disposition effect can be explained by loss aversion, but Hens and Vlcek (2006) and Barberis and Xiong (2006) have recently shown that loss aversion mostly does not predict the disposition effect. Muermann and Volkman (2007) have shown that regret theory can explain the disposition effect provided that  $Q(\cdot)$  is convex. Our findings of predominantly convex  $Q(\cdot)$  provide empirical support for their theoretical analysis and suggest that the disposition effect may indeed be caused by regret.

## **7. Discussion**

Our findings on utility were in line with other studies that used the tradeoff method in spite of the differences in experimental stimuli. This lends support to the robustness of utility measurements by the tradeoff method. Most subjects had concave utility. At the aggregate level utility was, however, close to linear. A possible heuristic that subjects could have applied in the elicitation of  $u(\cdot)$  by the tradeoff method would be to select a fixed difference between  $x_j$  and  $x_{j+1}$  and to choose option A if the difference was less than this value and B if it exceeded this value. Applying this heuristic consistently throughout the elicitation would result in linearity of  $u(\cdot)$ . There was little indication that subjects applied this heuristic: only two subjects displayed constant differences between successive elements of the standard sequence. The heuristic did not show up in the classification of the individual subjects either: even allowing for response error, few subjects were classified as having linear utility. It could of course be that subjects did not base their answers solely on this heuristic but that it nevertheless biased their responses in the direction of linearity of  $u(\cdot)$ . Abdellaoui et al. (forthcoming-a) compared the utilities elicited by tradeoff method

with those of other elicitation methods, which are less vulnerable to this heuristic, and found no significant differences.

In keeping with the predictions of regret theory,  $Q(\cdot)$  was found to be convex on  $[\frac{1}{5}, 1]$ . On  $[0, \frac{2}{5}]$   $Q(\cdot)$  was concave. One explanation for the initial concavity of  $Q(\cdot)$  is that regret aversion is not a universal phenomenon and that for small money amounts people do not feel regret (remember that the elicitation of  $Q(\cdot)$  on  $[0, \frac{2}{5}]$  involved the lowest money amounts). Regret is an emotion, which may only be triggered if the amounts at stake are sufficiently important. That previous studies found regret aversion for small money amounts does not challenge the above conjecture as the results from these studies were confounded by event-splitting effects.

There was little evidence in the data that the initial concavity of  $Q(\cdot)$  was caused by systematic biases in the elicited values of the elements of the standard sequence  $x_1, \dots, x_5$ . If  $x_1$  were overstated relative to  $x_2$  then this could explain the initial concavity of  $Q(\cdot)$ . For 24 subjects  $x_1 - 200$  exceeded  $x_2 - x_1$  and for 21 subjects  $x_2 - x_1$  exceeded  $x_1 - 200$ . The proportions are not significantly different ( $p = 0.766$ ). Of course, the modal preference of concave  $u(\cdot)$  would imply that more subjects had  $x_2 - x_1$  larger than  $x_1 - 200$ . The degree of concavity of  $u(\cdot)$  was, however, modest in our study and, hence, we do not believe that an overstatement of  $x_1$  relative to  $x_2$  is the main reason for the initial concavity of  $Q(\cdot)$ .

A more plausible explanation for the initial concavity of  $Q(\cdot)$  might be that other factors besides regret affected subjects' choices and that these factors tended to bias  $Q(\cdot)$  downwards on  $[\frac{1}{5}, 1]$ . Such a bias would exist if the elicited indifference probability  $p_{j+1}$ ,  $j = 1, \dots, 4$ , were too high. Recall that we elicited  $Q(\cdot)$  through a series of choices between  $(\text{€}200, p_{j+1}; \text{€}x_j)$  and  $(\text{€}x_{j+1}, p_{j+1}; \text{€}200)$ ,  $j = 1, \dots, 4$ . A plausible strategy to facilitate these choices was to cancel the common outcome  $\text{€}200$ , consistent with Kahneman and Tversky's (1979) isolation effect, and to reframe the decision problem as a choice between  $\text{€}x_{j+1}$  with probability  $p_{j+1}$  and  $\text{€}x_j$  with probability  $1 - p_{j+1}$ . Subjects could then have anchored on  $p_{j+1} = 0.50$  and subsequently have adjusted  $p_{j+1}$  downwards to account for the fact that  $x_{j+1}$  exceeded  $x_j$ . Psychological evidence indicates that such an anchoring and adjustment strategy leads to insufficient adjustment

(Hershey and Schoemaker 1985, Johnson and Schkade 1989), which in our case would lead to too high a value of  $p_{j+1}$  and, consequently, to too low a value of  $Q(\frac{j+1}{5})$ .<sup>12</sup>

Our study used forced choices: subjects always had to choose one of the options and indifference was not allowed. This was done to stimulate subjects to think hard about the choices and to avoid “lazy responses”: subjects might just state indifference because it is an easy option. The problem with the use of forced choices is that if subjects are truly indifferent they still have to choose one of the options. If they break ties in a random manner then there is no problem. If they use a specific rule to break ties, however, then the use of forced choices may have biased the results. It turned out that selecting a specific rule for breaking ties had no effect on  $u(\cdot)$ . For the elicitation of  $Q(\cdot)$ , however, it might have had an effect. Breaking ties by consistently choosing option A (B) somewhat biased the elicitation of  $Q(\cdot)$  in the direction of concavity (convexity). If subjects applied a consistent rule to break ties, then it is more plausible that they would do so by choosing A because A was the less risky option and most subjects were risk averse. Hence, we are inclined to believe that if the use of forced choices has biased the results then this bias is in the direction of concavity of  $Q(\cdot)$  and our finding of predominantly convex  $Q(\cdot)$  cannot be attributed to the use of forced choices.

Regret theory assumes that people evaluate probabilities in a linear manner. Many studies have shown, however, that people are not well able to handle probabilities and evaluate them in a nonlinear manner. These studies have also shown that there is little probability transformation for the median probabilities that we elicited, which were all between 0.40 and 0.43. Hence, it seems implausible that probability transformation alone can explain the observed shape of  $Q(\cdot)$ . Indeed, prospect theory was not entirely consistent with our data either. Because our method for eliciting utility is also valid under prospect theory, the median indifference  $(x_2, 0.40, \text{€}200) \sim (\text{€}200, 0.40, x_1)$ , for instance, implied under prospect theory that  $w(0.60)$  was twice as large as  $w(0.40)$ . Previous evidence has indicated, however, that  $w(0.60)$  is only about 1.3 times as large as  $w(0.40)$  (Tversky and Kahneman 1992, Gonzalez and Wu 1999, Abdellaoui 2000, Bleichrodt and Pinto 2000).

---

<sup>12</sup> This evidence suggests that anchoring and insufficient adjustment affects matching tasks in particular. Delqu e 1997 showed, however, that choice-based elicitation can also be vulnerable to some of the biases that affect matching.

A troubling finding for regret theory is that one of the consistency tests, the test for  $Q(2/5)$ , led to significantly different results. Recall that  $Q(2/5)$  was elicited in the consistency test by determining indifference between  $(\text{€}x_3, p_{c1}; \text{€}x_2)$  and  $(\text{€}x_4, p_{c1}; \text{€}200)$ . We observed that the indifference value of  $p_{c1}$  was significantly higher than the value that we would expect based on the elicitation of  $Q(\cdot)$ . One explanation for this discrepancy might be reference-dependence. Consider our median subject for whom  $p_{c1}$  was determined through choices between  $(\text{€}810, p_{c1}; \text{€}610)$  and  $(\text{€}1000, p_{c1}; \text{€}200)$ . Because the first option gives the certainty of receiving at least  $\text{€}610$ , subjects may have taken  $\text{€}610$  as their reference point and have evaluated the other amounts as gains and losses relative to this reference point. Loss aversion can then explain the high value of  $p_{c1}$  that we observed. Such an explanation of our data would be in line with prospect theory. Taken together, our findings seem to confirm Loomes and Sugden's (1982) conclusion that "we do not believe that [...] utility and regret are the only factors that influence behaviour under uncertainty, but just that these factors seem particularly significant." (p. 819).

## **8. Conclusion**

This paper is the first to make regret theory, a popular and influential theory of decision under uncertainty that was originally proposed by Bell (1982) and Loomes and Sugden (1982), quantitatively observable. We have shown that such a quantitative measurement is feasible and can be performed by familiar measurement tools, which will hopefully stimulate future applications of regret theory in economics and decision analysis. Our method controlled for event-splitting effects, which have been found to distort earlier tests of regret theory. Our findings were to a large extent consistent with the predictions of regret theory although some deviations were observed. These deviations underline that although utility and regret seem to be important in decision under uncertainty, other factors are relevant as well. In the context of our experiment, we conjecture that anchoring and insufficient adjustment and reference-dependence may be two such factors.

## Appendix A: Method for Eliciting the Indifference Values

In the measurement of  $u(\cdot)$ ,  $x_{j+1}$  was elicited through choices between  $A = (\text{€}x_j, \frac{1}{3}; \text{€}100)$  and  $B = (\text{€}x_{j+1}, \frac{1}{3}; \text{€}40)$ ,  $j = 0, \dots, 4$ . The stimulus  $x_{j+1}$  was initially set so that A and B had equal expected value, i.e.  $x_{j+1} = x_j + 120$ . There were two possible scenarios:

(i) If A was chosen we added 100 to  $x_{j+1}$  until B was chosen. We then subtracted 40 from  $x_{j+1}$ . If A was chosen 20 was added. If B was chosen we subtracted 20 until A was chosen. The program ended in case A was still chosen when  $x_{j+1} = x_j + 820$ .

(ii) If B was chosen we subtracted 100 from  $x_{j+1}$ . If B was still chosen we subtracted 20, so that A dominated B. If subjects still chose B a warning appeared: “You chose B even though A yields at least as good an outcome as B under both events. Are you sure you prefer B?” If the subject answered “yes”, the experiment ended. If the subject answered “no”, the indifference value was set equal to  $x_j + 10$ . If A was chosen 40 was added. If subsequently B was chosen 20 was subtracted. If subsequently A was chosen 20 was added until B was chosen.

The recorded indifference value was the midpoint between the lowest value of  $x_{j+1}$  for which B was chosen and the highest value of  $x_{j+1}$  for which A was chosen. Table A1 gives an example of the procedure for the elicitation of  $x_1$  through comparisons between  $A = (\text{€}200, \frac{1}{3}; \text{€}100)$  and  $B = (\text{€}x_1, \frac{1}{3}; \text{€}40)$ . In this example, the recorded indifference value was the midpoint of 380 and 360, that is, 370.

**Table A1: Example of the elicitation of  $x_1$**

Iteration	$x_1$	Choice
1	320	A
2	420	B
3	380	B
4	360	A

The procedure in the second stage was largely similar. We elicited the value of  $p_{j+1}$  for which indifference held between  $A = (\text{€}200, p_{j+1}; \text{€}x_j)$  and  $B = (\text{€}x_{j+1}, p_{j+1}; \text{€}200)$ ,  $j = 1, \dots, 4$ . We started by

selecting  $p_{j+1}$  so that A and B had equal expected value. Each time A was chosen we selected the next value of  $p_{j+1}$  as the midpoint between the value of  $p_{j+1}$  in the last iteration and the smallest probability used in the previous iterations that exceeded the probability in the last iteration or, if such a probability did not exist, as the midpoint between the value of  $p_{j+1}$  in the last iteration and 1. Each time B was chosen we selected the next value of  $p_{j+1}$  as the midpoint between the value of  $p_{j+1}$  in the last iteration and the largest probability used in the previous iterations that was smaller than the probability in the last iteration or, if such a probability did not exist, as the midpoint between the value of  $p_{j+1}$  in the last iteration and 0. If the probability was not a multiple of 0.01 we rounded it upwards to the next multiple of 0.01. Table A2 illustrates the procedure for  $x_{j+1} = €650$  and  $x_j = €390$ . The recorded indifference value is 0.22 in the example in the table.

**Table A2: Example of the elicitation of  $p_{j+1}$  when  $x_{j+1} = 650$  and  $x_j = 390$**

Iteration	$p_{j+1}$	Choice
1	0.30	B
2	0.15	A
3	0.23	B
4	0.19	A
5	0.21	A

## Appendix B: Overview of the Individual Results

	$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$p_1$	$p_2$	$p_3$	$p_4$	$p_5$	Pow. coeff. for $u(\cdot)$	Exp. coeff. for Q
1	490	620	730	860	990	0.50	0.47	0.46	0.47	0.48	1.729	-3.314
2	450	840	1150	1480	1630	0.50	0.40	0.42	0.42	0.44	1.238	-0.064
3	350	560	750	900	1170	0.50	0.15	0.22	0.32	0.44	0.746	1.795
4	290	360	470	540	630	0.50	0.42	0.39	0.41	0.41	0.997	0.806
5	590	800	1010	1300	1490	0.50	0.49	0.49	0.48	0.48	1.346	-5.959
6	630	940	1470	1940	2490	0.50	0.43	0.43	0.41	0.42	0.799	0.345
7	650	900	1010	1180	1370	0.50	0.29	0.22	0.42	0.37	1.852	2.114
8	390	460	590	720	850	0.50	0.41	0.39	0.40	0.43	1.165	0.567
9	390	580	670	780	970	0.50	0.40	0.47	0.47	0.40	1.204	-0.421
10	410	580	710	1000	1190	0.50	0.42	0.40	0.41	0.47	0.855	-0.436
11	670	980	1690	2200	2990	0.50	0.37	0.31	0.21	0.20	0.685	6.829
12	370	680	1010	1320	1530	0.50	0.21	0.09	0.23	0.14	0.888	8.870
13	530	620	730	860	990	0.50	0.38	0.43	0.47	0.45	1.821	-1.154
14	390	400	790	900	1110	0.50	0.51	0.38	0.48	0.50	0.731	-2.953
15	670	940	1550	2140	2650	0.50	0.49	0.50	0.47	0.50	0.796	-8.359
16	230	300	490	560	630	0.50	0.31	0.39	0.42	0.40	0.466	1.022
17	390	600	850	1100	1390	0.50	0.47	0.44	0.44	0.45	0.718	-1.063
18	490	1060	1490	1600	1770	0.50	0.25	0.19	0.47	0.23	1.884	4.637
19	530	860	990	1180	1330	0.50	0.48	0.48	0.47	0.45	1.769	-2.671
20	210	840	1270	1640	2370	0.50	0.01	0.39	0.45	0.42	0.524	0.582
21	550	800	1010	1520	1830	0.50	0.43	0.37	0.40	0.42	0.837	0.826
22	490	700	1010	1340	1450	0.50	0.43	0.44	0.41	0.47	1.19	-0.819

23	490	600	810	920	1190	0.50	0.48	0.50	0.48	0.41	1.081	-1.874
24	330	460	590	720	850	0.50	0.36	0.42	0.45	0.46	1.000	-0.861
25	290	380	510	640	770	0.50	0.32	0.39	0.43	0.43	0.574	0.339
26	310	500	730	900	1030	0.50	0.49	0.35	0.32	0.37	0.939	2.476
27	570	900	1230	1600	1890	0.50	0.38	0.47	0.44	0.47	1.082	-1.549
28	350	500	630	800	1110	0.50	0.40	0.40	0.36	0.34	0.478	2.600
29	430	500	590	660	750	0.50	0.30	0.43	0.48	0.34	2.091	1.285
30	310	680	1050	1580	2210	0.50	0.38	0.38	0.48	0.51	0.305	-2.327
31	370	540	1170	1580	1990	0.50	0.30	0.39	0.40	0.46	0.420	0.178
32	310	480	910	1280	1530	0.50	0.44	0.44	0.42	0.42	0.39	0.101
33	330	420	550	680	810	0.50	0.39	0.29	0.44	0.46	0.868	0.107
34	390	580	790	1000	1310	0.50	0.40	0.45	0.44	0.41	0.685	0.039
35	210	300	430	500	510	0.50	0.05	0.24	0.40	0.51	1.195	0.123
36	390	500	650	860	1030	0.50	0.41	0.39	0.43	0.43	0.869	0.186
37	390	700	990	1360	1790	0.50	0.28	0.41	0.33	0.41	0.579	1.69
38	390	720	950	1100	1270	0.50	0.30	0.38	0.37	0.37	1.390	2.100
39	490	620	810	940	1070	0.50	0.47	0.43	0.44	0.45	1.635	-0.943
40	370	760	950	1060	1170	0.50	0.36	0.21	0.44	0.44	1.978	0.802
41	570	740	1130	1300	1790	0.50	0.43	0.43	0.45	0.43	0.834	-0.485
42	610	780	970	1180	1490	0.50	0.47	0.49	0.50	0.45	1.190	-4.292
43	310	400	630	1000	1530	0.50	0.31	0.35	0.38	0.40	-0.124	1.583
44	290	460	690	900	1070	0.50	0.48	0.42	0.40	0.47	0.589	-0.634
45	290	380	490	660	830	0.50	0.44	0.40	0.41	0.39	0.299	1.100
46	430	660	770	940	1250	0.50	0.36	0.44	0.46	0.38	0.893	0.554
47	490	860	1190	1560	1830	0.50	0.38	0.42	0.41	0.43	1.019	0.315
48	330	460	590	720	850	0.50	0.36	0.39	0.42	0.46	1.000	-0.172
49	450	560	810	1240	1530	0.50	0.47	0.43	0.43	0.41	0.571	0.162
50	510	680	830	1000	1170	0.50	0.41	0.42	0.43	0.44	1.471	-0.237
51	370	620	890	1420	1850	0.50	0.32	0.40	0.33	0.40	0.373	1.861
52	290	440	550	700	850	0.50	0.35	0.43	0.43	0.43	0.676	0.034
53	410	640	810	940	1070	0.50	0.31	0.31	0.31	0.22	1.552	5.82
54	350	520	670	840	1010	0.50	0.39	0.40	0.41	0.41	0.921	0.799
55	370	620	870	1080	1310	0.50	0.36	0.41	0.42	0.43	0.898	0.284
56	290	380	530	660	790	0.50	0.40	0.38	0.36	0.35	0.584	2.477

## References

- Abdellaoui, M. 2000. Parameter-free elicitation of utility and probability weighting functions. *Management Science* **46** 1497-1512.
- Abdellaoui, M., C. Barrios, P. P. Wakker. forthcoming-a. Reconciling introspective utility with revealed preference: Experimental arguments based on prospect theory. *Journal of Econometrics*.
- Abdellaoui, M., H. Bleichrodt, C. Paraschiv. forthcoming-b. Measuring loss aversion under prospect theory: A parameter-free approach. *Management Science*.
- Abdellaoui, M., F. Vossman, M. Weber. 2005. Choice-based elicitation and decomposition of decision weights for gains and losses under uncertainty. *Management Science* **51** 1384-1399.
- Anand, P. 1987. Are the preference axioms really rational? *Theory and Decision* **23** 189-214.
- Barberis, N., M. Huang, R. H. Thaler. 2006. Individual preferences, monetary gambles, and stock market participation: A case for narrow framing. *American Economic Review* **96** 1069-1090.
- Barberis, N., W. Xiong. 2006. What drives the disposition effect? An analysis of a long-standing preference-based explanation. *Working Paper, Yale University and Princeton University*.
- Beattie, J., G. Loomes. 1997. The impact of incentives upon risky choice experiments. *Journal of Risk and Uncertainty* **14** 155-168.
- Bell, D. E. 1982. Regret in decision making under uncertainty. *Operations Research* **30** 961-981.
- Bell, D. E. 1983. Risk premiums for decision regret. *Management Science* **29** 1156-1166.

- Bell, D. E. 1985. Putting a premium on regret: Reply. *Management Science* **31** 117-120.
- Bleichrodt, H., J. L. Pinto. 2000. A parameter-free elicitation of the probability weighting function in medical decision analysis. *Management Science* **46** 1485-1496.
- Bostic, R., R. J. Herrnstein, R. D. Luce. 1990. The effect on the preference reversal of using choice indifferences. *Journal of Economic Behavior and Organization* **13** 193-212.
- Braun, M., A. Muermann. 2004. The impact of regret on the demand for insurance. *Journal of Risk and Insurance* **71** 737-767.
- Delquié, P. 1997. 'Bi-matching': A new preference assessment method to reduce compatibility effects. *Management Science* **43** 640-658.
- Etchart-Vincent, N. 2004. Is probability weighting sensitive to the magnitude of consequences? An experimental investigation on losses. *Journal of Risk and Uncertainty* **28** 217-235.
- Fechner. 1860/1966. *Elements of psychophysics*. Holt, Rinehart, and Winston, New York.
- Fennema, H., M. van Assen. 1999. Measuring the utility of losses by means of the trade-off method. *Journal of Risk and Uncertainty* **17** 277-295.
- Fishburn, P. C. 1982. Non-transitive measurable utility. *Journal of Mathematical Psychology* **26** 31-67.
- Fishburn, P. C. 1991. Nontransitive preferences in decision theory. *Journal of Risk and Uncertainty* **4** 113-134.
- Genesove, D., C. J. Mayer. 2001. Loss aversion and seller behavior: Evidence from the housing market. *Quarterly Journal of Economics* **116** 1233-1260.
- Goeree, J., C. A. Holt, T. Palfrey. 2003. Risk averse behavior in generalized matching pennies games. *Games and Economic Behavior* **45** 97-113.
- Gollier, C., B. Salanié. 2006. Individual decisions under risk, risk sharing and asset prices with regret. *Working Paper, University of Toulouse and INSEE*.
- Gonzalez, R., G. Wu. 1999. On the form of the probability weighting function. *Cognitive Psychology* **38** 129-166.
- Grinblatt, M. S., M. Keloharju. 2001. What makes investors trade? *Journal of Finance* **56** 589-616.
- Harless, D., C. F. Camerer. 1994. The predictive utility of generalized expected utility theories. *Econometrica* **62** 1251-1289.
- Heath, C., S. Huddart, M. Lang. 1999. Psychological factors and stock option exercise. *Quarterly Journal of Economics* **114** 601-627.
- Hens, T., M. Vlcek. 2006. Does prospect theory explain the disposition effect? *Working Paper, University of Zurich*.
- Hershey, J. C., P. J. H. Schoemaker. 1985. Probability versus certainty equivalence methods in utility measurement: Are they equivalent? *Management Science* **31** 1213-1231.
- Hey, J. D., C. Orme. 1994. Investigating generalizations of expected utility theory using experimental data. *Econometrica* **62** 1291-1326.

- Humphrey, S. J. 1995. Regret aversion or event-splitting effects? More evidence under risk and uncertainty. *Journal of Risk and Uncertainty* **11** 263-274.
- Johnson, E. J., D. A. Schkade. 1989. Bias in utility assessments: Further evidence and explanations. *Management Science* **35** 406-424.
- Kahneman, D., A. Tversky. 1979. Prospect theory: An analysis of decision under risk. *Econometrica* **47** 263-291.
- Krantz, D. H., R. D. Luce, P. Suppes, A. Tversky. 1971. *Foundations of measurement, vol. 1*. Academic Press, New York.
- Larrick, R. P. 1993. Motivational factors in decision theories: The role of self-protection. *Psychological Bulletin* **113** 440-450.
- Larrick, R. P., T. L. Boles. 1995. Avoiding regret in decisions with feedback: A negotiation example. *Organizational Behavior and Human Decision Processes* **63** 87-97.
- Loomes, G. 2005. Modelling the stochastic component of behaviour in experiments: Some issues for the interpretation of data. *Experimental Economics* **8** 301-323.
- Loomes, G., C. Starmer, R. Sugden. 1991. Observing violations of transitivity by experimental methods. *Econometrica* **59** 425-439.
- Loomes, G., R. Sugden. 1982. Regret theory: An alternative theory of rational choice. *Economic Journal* **92** 805-824.
- Loomes, G., R. Sugden. 1987. Some implications of a more general form of regret theory. *Journal of Economic Theory* **41** 270-287.
- Loomes, G., R. Sugden. 1995. Incorporating a stochastic element into decision theories. *European Economic Review* **39** 641-648.
- Loomes, G., R. Sugden. 1998. Testing different stochastic specifications of risky choice. *Economica* **65** 581-598.
- McKelvey, R., T. Palfrey. 1995. Quantal response equilibria for normal form games. *Games and Economic Behavior* **10** 6-38.
- Michenaud, S., B. Solnik. 2006. Applying regret theory to investment choices: Currency hedging decisions. *Working Paper, HEC-School of Management*.
- Muermann, A., O. S. Mitchell, J. M. Volkman. 2006. Regret, portfolio choice, and guarantees in defined contribution schemes. *Insurance: Mathematics and Economics* **39** 219-229.
- Muermann, A., J. M. Volkman. 2007. Regret, pride, and the disposition effect. *Working Paper, University of Pennsylvania*.
- Odean, T. 1998. Are investors reluctant to realize their losses? *Journal of Finance* **53** 1775-1798.
- Rothschild, M., J. E. Stiglitz. 1970. Increasing risk I: A definition. *Journal of Economic Theory* **2** 225-243.
- Schunk, D., C. Betsch. 2006. Explaining heterogeneity in utility functions by individual differences in decision modes. *Journal of Economic Psychology* **27** 386-401.

- Shefrin, H. M., M. Statman. 1985. The disposition to sell winners too early and ride losers too long. *Journal of Finance* **40** 777-790.
- Smith, R. D. 1996. Is regret theory an alternative basis for estimating the value of health care interventions? *Health Policy* **37** 105-115.
- Starmer, C. 1992. Testing new theories of choice under uncertainty using the common consequence effect. *Review of Economic Studies* **59** 813-830.
- Starmer, C. 1999. Cycling with rules of thumb: An experimental test for a new form of non-transitive behavior. *Theory and Decision* **46** 141-158.
- Starmer, C., R. Sugden. 1991. Does the random-lottery incentive system elicit true preferences? An experimental investigation. *American Economic Review* **81** 971-978.
- Tversky, A., D. Kahneman. 1992. Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty* **5** 297-323.
- Wakker, P. P. 2006. The power (crra) utility family graphically explained. *Working Paper, Erasmus University Rotterdam*.
- Wakker, P. P., D. Deneffe. 1996. Eliciting von Neumann-Morgenstern utilities when probabilities are distorted or unknown. *Management Science* **42** 1131-1150.
- Yaniv, G. 2000. Withholding information from cancer patients as a physician's decision under risk. *Medical Decision Making* **20** 216-227.
- Zeelenberg, M. 1999. Anticipated regret, expected feedback and behavioral decision making. *Journal of Behavioral Decision Making* **12** 93-106.
- Zeelenberg, M., J. Beattie, J. van der Pligt, N. K. de Vries. 1996. Consequences of regret aversion: Effects of expected feedback on risky decision making. *Organizational Behavior and Human Decision Processes* **65** 148-158.