Estimating risk attitudes using lotteries; a large sample approach

Bas Donkers

Bertrand Melenberg

Arthur van Soest[1]

Abstract

Attitudes towards risk play a major role in many economic decisions. In empirical studies one quite often assumes that attitudes towards risk do not vary across individuals. This papers questions this assumption and analyses which factors influence an individual's risk attitude. Based on questions on lotteries in a large household survey we semiparametrically estimate an index for risk aversion. We only make weak assumptions about the underlying decision process, and our estimation method allows for generalisations of expected utility. We find strong links between risk aversion and gender, education level, and income of the individual. We also estimate a structural model based on Cumulative Prospect Theory and find that the value function depends on an index that is very similar to the index of risk aversion. Expected utility is strongly rejected and the probability weighting function varies significantly with gender, age, and income of the individual.

Keywords: Risk aversion, non-expected utility, semiparametric estimation

---

[1]We thank Peter Wakker and Arie Kapteyn for helpful comments. Financial support for the first author has been provided by the Netherlands Organization for Scientific Research (NWO).

Correspondence to: Bas Donkers, Tilburg University, Department of Econometrics and CentER, P.O. Box 90153, 5000 LE Tilburg, The Netherlands (e-mail: Donkers@kub.nl)

1 Introduction

Attitudes towards risk are important in many economic decisions. In empirical studies of economic behavior, however, direct information about attitudes towards risk is hardly ever available. This paper uses a large Dutch household survey that contains both direct information on respondents' attitudes towards risk and a lot of background information on the respondents. We use these data to investigate whether attitudes towards risk vary with other observed characteristics fo the respondents, such as age and income. Whether and how an individual's attitude towards risk varies with observed characteristics can be helpful in empirical studies where this type of information is not present, but the background characteristics are observed.

Our inference on attitudes towards risk is based upon a set of eight questions on lotteries that are present in the data. In five of these questions the respondents have to make a choice between two lotteries. The remaining three questions are probability equivalence questions. Here the respondents have to state the minimum probability of winning a given prize, which would make them indifferent between such a lottery and a given amount of money. Both types of questions have a risky (high variance) and a safe (low or zero variance) option. We use these data to distinguish between more and less risk averse individuals.

To see how an individual's attitude towards risk relates to other observed characteristics we start with a very general semiparametric model. We do not use any economic or psychological theory, but we only impose a single index restriction and a monotonicity condition, such that the index represents the individual's risk aversion. The estimation results show a significant relationship between risk aversion and age, gender, education level, and income.

The semiparametric model is too general to permit a clear cut interpretation of the consequences of differences in attitudes towards risk. Therefore, we set up a structural model for the individual's decision process. Expected utility theory seems a good starting point in analyzing decisions under uncertainty. However, within the experimental psychology literature considerable evidence is reported against the validity of expected utility when individuals answer questions on lotteries, see, for example, Kahneman, Slovic and Tversky (1982), or Machina (1987). Instead of the expected utility framework we will use Cumulative Prospect Theory as developed by Kahneman and Tversky (1979) and Tversky and Kahneman (1992). Expected utility can be seen as a special case of Cumulative Prospect Theory. The index for risk aversion we find with the structural model is quite similar to the one we obtained with the semiparametric model. This gives us an interpretation of the semiparametric estimation results.

The approach we take is possible since the data contain the questions related to risk attitudes as well as many background variables for almost 4,000 individuals. This contrasts with the datasets that have been used until now to derive measures for individuals' attitudes towards risk. In the experimental psychology and economics literature the datasets are, in general, rather small, consisting of no more than 200 individuals, and contain hardly any background information. The respondents in these studies are also very often students and the results are thus not representative of the population of interest. The presence of small datasets is illustrated by the fact that Harless and Camerer (1994) merge a total of 23 datasets to obtain nearly 8,000 choices, where at least 3 choices are made by each individual. Our results will be based on more than 20,000 choices. In the economics literature an indirect measure of risk aversion is sometimes derived from observed behavior, but these results are quite sensitive to many real life aspects that are unrelated to risk aversion. Example of this line of research are Pålsson (1996), who estimates risk aversion from portfolio choices, and Guiso et al. (1992), who derive a measure of risk aversion from savings data. A final branch in the literature uses large datasets from, for instance, TV shows or bets on horse races, but these datasets address very specific populations and contain no background information on the individuals making the decisions (see, for example, Beetsma and Schotman (1997) and Jullien and Salanie (1997)).

An exception is Hartog et al. (1997), who use a dataset which contains almost 2,000 individuals and a lot of background information, but direct information on risk attitudes is provided by only a single question.

Information on how risk aversion varies across individuals can be useful in predicting savings or stock holdings of an individual or household, since risk aversion plays an important role in these decisions. However, the existing empirical literature on modelling savings and portfolio choices focuses mainly on the effect of income risk and makes restrictive assumptions about the individuals' attitudes towards risk. For example, Lusardi (1997) estimates a single coefficient for risk aversion, which is constant across all individuals, while Guiso et al. (1992) allow risk aversion to depend only on lifetime resources. The present paper shows that attitudes towards risk also vary with other individual characteristics. Our analysis indicates which variables should be used to model attitudes towards risk in empirical applications, where no direct information on risk attitudes is available.

The remainder of the paper is structured as follows. We start with a detailed description of the data in Section 2. Section 3 presents the reduced form model and the semiparametric estimation techniques we will use, while Section 4 presents the semiparametric estimation results. Section 5 discusses the individual's decision making process, where we pay special attention to Cumulative Prospect Theory. Section 6 presents the structural model based on Cumulative Prospect Theory and its estimation results. Section 7 concludes.

2 Data

The data we use come from the first wave of the VSB-Panel data, drawn in 1993, and consist of 2780 households, divided into two panels. One is designed to be representative of the Dutch population, the other is a random sample of the households in the upper 10% of the income distribution in The Netherlands. All households participating have been provided with a personal computer and answer the survey questions directly on their PC; no personal interviews are held. The VSB-panel is a rich source of data, including information on household composition, income, assets and psychological concepts. A detailed description of the data can be found in Nyhus (1996). The total numbers of households in the representative and high income panel answering the relevant psychological questionnaire are 1463 and 783, with a total number of individual respondents of 2297 and 1652, respectively.

This psychological questionnaire contains a set of questions on lotteries. This set of questions consists of two types of questions (see Appendix A.1 for the precise wording of the questions):

1. The first type of question deals with choices between two lotteries. Each time the respondent is offered two lotteries, each one with two possible outcomes with given probabilities,[2] and the respondent has to state which of the two lotteries he or she prefers. It is mentioned that there do not exist right or wrong answers to these questions. We refer to these questions as choice questions.

   Five questions of this type are asked with varying outcomes and probabilities.

2. The second type of question deals with the imaginary situation where a certain amount of money has been won and the individual has the opportunity to buy a lottery ticket with this money. This lottery ticket has a single prize of Dfl 20,000[3] and the question is how large the probability of winning this Dfl 20,000 has to be at least, to make the respondent willing to exchange the money for the lottery ticket. The amount of money that can be won and is exchanged for the lottery ticket is varied over the questions. We refer to these questions as probability equivalence questions.

   Three questions of this type are asked.

The answers to the questions of the first type will be referred to by CH^1, CH^2, $\ldots, CH^5$ and are depicted in Table 1. We call the low variance lottery the safest

[2] In one case one alternative is winning zero with probability one.

[3] Dfl 1 was approximately US$ 0.50 by the end of 1993

option and the high variance lottery the riskiest option. A value 1 corresponds to the choice of the risky option, while 0 indicates that the safe option is chosen. Individuals opting for the safe lottery are called more risk averse than individuals who choose the risky option.

The answers to the questions of the second type, the probability equivalence questions, will be referred to by PE^1, PE^2, and PE^3. The answers indicate the probability (in %) of winning the prize of Dfl 20,000 for which the individual is indifferent between the lottery and an amount of money for sure of, respectively, Dfl 200 (PE^1), Dfl 1000 (PE^2), and Dfl 5000 (PE^3). The variables range from 0% to 100%. A higher probability of winning the prize implies a more attractive lottery. A more risk averse individual will thus give higher answers. The fact that a higher probability of winning corresponds to a more attractive lottery also implies a logical consistency requirement that $PE^1 < PE^2 < PE^3$, if marginal utility of money is positive (more is better). Otherwise individuals would prefer a stochastically dominated alternative. By now it is well known that probability equivalence questions result in overestimation of the level of risk aversion, due to, for example, response mode bias. In the questions we analyze individuals who have to give up money to participate, which might strengthen this bias.

In total we have 3949 individuals in our sample if we use both the representative panel and the high income panel. In the final analysis we will condition upon income, so there will be no effect of the overrepresentation of high income households. Sample means and other unconditional statistics will be reported for the representative panel only, so the numbers we present are representative of the Dutch population. For 491 respondents we miss important demographic information such as age or education, but mostly individual income, leaving us with 3458 observations. Furthermore, there are 865 individuals giving the answer "Don't know" to at least one of the probability equivalence questions. Most of them did not answer any question. This might be caused by lack of interest in this type of questions, but it can also be the case that these questions are rather difficult (see Warneryd (1996) for a discussion of this problem). We do not use observations with one or more missing values to the probability equivalence questions, since these respondents might not really understand the questions. The sample we use for estimation consists of 2593 individuals for whom we observe both tha answers to the questions on the lotteries and the individual characteristics we want to use in explaining the individuals risk attitudes. These include 237 respondents who gave an inconsistent set of answers, satisfying either $PE^1 > PE^2$, $PE^2 > PE^3$ or $PE^1 > PE^3$.

The fraction of respondents choosing the riskiest option in the choice questions are presented in Table 1. The table shows that the number of individuals choosing

Table 1: Descriptive statistics for the choice questions, representative panel only.

Question	fraction choosing riskiest lottery	Safest		Riskiest	
		μ	σ	μ	σ
CH^1 (1000;1) vs. (2000;0.5)	0.21	1000	0	1000	1000
CH^2 (30;1) vs. (45;0.8)	0.40	30	0	36	18
CH^3 (100;0.25) vs. (130;0.20)	0.49	25	42	26	52
CH^4 (3000;0.02) vs. (6000;0.01)	0.56	60	420	60	597
CH^5 (0;1) vs. (1500;0.5,-1000;0.5)	0.12	0	0	250	1250

Note: $(x;p)$ denotes the lottery paying x with probability p and zero otherwise, while $(x;p,y;q)$ denotes the lottery paying x with probability p and y with probability q.

the riskiest lottery varies considerably across the questions. This is largely due to the difference in expected value, μ, between the two lotteries relative to the difference in risk, σ, taken. For CH^1 and CH^4, there is no reward for the extra risk taken, i.e., the expected value of the two lotteries is the same. Respondents choosing the riskiest option in one of these two questions show risk loving behavior. Note, however, that some non-expected utility theories are able to explain this behavior, even if the marginal utility of money is decreasing, which would be equivalent with the regular concept of risk aversion under expected utility as it is defined by Pratt (1964) and

Arrow (1965). Aspects of these theories that can be relevant are the certainty effect in CH^1 and CH^2. For CH^4 subproportionality can be important, while for CH^5 it is loss aversion.

The mean and median of the probability equivalence questions can be found in Table 2. The mean of the answers to PE^1 is a 34% chance of winning Dfl 20,000, while the median answer to this question is 25%. The other columns have to be read in a similar way. There is a clear pattern of increasing answers if we go from PE^1 to PE^3, but there is also substantial variation across respondents for each question.

Table 2: Descriptive statistics for the probability equivalence questions

	PE^1	PE^2	PE^3
	(200;1) vs. (20,000;p)	(1000;1) vs. (20,000;p)	(5000;1) vs. (20,000;p)
Mean	34.0%	45.6%	62.3%
Median	25%	50%	60%

3 A semiparametric model for risk attitudes

Many papers have estimated attitudes towards risk using specific functional forms to represent preferences, see, for example, Lattimore, Baker, and Witte (1992), Tversky and Kahneman (1992), and Beetsma and Schotman (1997). Exceptions are Wakker and Deneffe (1996) and Abdellaoui (1998).

In this section we will not specify any functional forms. For the choice questions we assume that $E\{1 - CH^q|x\} = P\{\text{safest choice is chosen in question } q|x\} = G^q(x'\beta^q)$, with x a vector of observed characteristics such as age and income. β^q is a parameter vector that has to be estimated. The function $G^q(.)$ is unknown, but assumed to be increasing. A higher value of $x'\beta^q$, the individual specific index, now implies a higher probability for the safe option being chosen and thus more risk aversion.

For the probability equivalence questions we make a similar assumption, which is that $E\{PE^q|x\} = F^q(x'\beta^q)$. x denotes a vector of observed characteristics and β^q a vector of parameters. Teh function $F^q(.)$ is not known, but assumed to be increasing. Higher values og $x'\beta^q$ imply, on average, a higher answer to the probability equivalence question. As was the case for the choice questions, a higher value of $x'\beta^q$ implies more risk aversion. Therefore we will refer to this index as a measure of risk aversion.

We do not assume that the index is the same for each question. One of the interesting questions now is whether the indices for each of the questions are the same. When we present the estimation results, we will also present the results of some tests on equivalence of the indiced. If, for example, loss aversion is stronger for one group of individuals and small probabilities are more overweighted by another group, we could obtain different estimates for CH^4 and CH^5. This then indicates that we are not able to model the respondents' behavior towards all the questions with a single index.

With these assumptions the scale of β^q is not identified and we normalize the component of β^q that relates to the individual's gender, say the first component, such that $|\beta_1^q| = 1$. The sign of β_1^q and thus whether females are more or less risk averse than males is identified. These assumptions and some technical regularity conditions are sufficient to obtain a consistent estimator for β^q for each question separately, using the rank estimator proposed by Cavanagh and Sherman (1998), which is an extension of the maximum rank correlation estimator of Han (1987). One practical problem with this estimator is that the objective function does not behave very well with the data we have, which is due to properties of the objective function in small samples. We solved this problem by replacing the objective function with a smoothed version, following the idea of Horowitz (1992), which, under appropriate regularity assumptions, does not change the asymptotic properties of the estimator. For the probability equivalence questions we can use this initial estimator and the method proposed by Delecroix, Härdle, and Hristache (1997) to obtain asymptotic efficiency by a one step estimator based on the initial estimate. For the choice questions we use the same approach, where the efficient estimator is based on the ideas in Klein and Spady (1993). Technical details about the estimators are given in Appendix B.

With the method described above, we obtain efficient estimates of β^q for each question. Since we are interested in a single measure of risk attitude per person and not in a measure of risk attitude per person for each question, we will test whether the estimated coefficients for the different questions are actually the same and thus whether there exists a unique measure of attitude towards risk for the questions we have. To combine the estimates from the different questions we use minimum distance (see Lee (1996), for example). Here we take into account the fact that we observe the same individual more than once and the estimates for the different questions are not independent.

4 Estimation results for the semiparametric model

This section presents the estimation results for the semiparametric model defined in the previous section. Tables 3 and 4 present the estimates for β and the corresponding standard errors for each of the questions separately. As explanatory variables in our model we use a dummy variable for gender, age, the logarithm of income, and education level measured on a scale from 1 to 5. For some of the respondents we did not observe their personal income, or it was zero. For these respondents Log(Income) was set to zero and a dummy, Dinczero, was included to correct for this. Some descriptive statistics of the explanatory variables are given in Appendix A.2.

The estimates are all significant, except for the effect of income on CH^1 and the effect of education on CH^5. The estimates for the three probability equivalence questions are very similar, even though the estimates are computed completely independent from each other. Also the signs of the estimates for CH^2, CH^3, and CH^4

Table 3: Estimation results for β for the choice questions. Standard errors are in parentheses.

	CH^1	CH^2	CH^3	CH^4	CH^5
Female	1 (——)	1 (——)	1 (——)	1 (——)	1 (——)
Age	0.032 (0.004)	0.031 (0.003)	0.031 (0.004)	0.214 (0.022)	0.031 (0.008)
Log(Income)	-0.080 (0.046)	-0.364 (0.048)	-0.171 (0.043)	-1.127 (0.134)	0.125 (0.058)
Education	0.088 (0.026)	-0.280 (0.030)	-0.088 (0.022)	-0.272 (0.041)	0.063 (0.037)
Dinczero	-0.578 (0.477)	-3.530 (0.488)	-1.930 (0.449)	-8.419 (1.056)	1.232 (0.551)

Table 4: Estimation results for β for the probability equivalence questions. Standard errors in parentheses.

	PE^1	PE^2	PE^3
Female	1 (——)	1 (——)	1 (——)
Age	0.029 (0.001)	0.031 (0.001)	0.032 (0.001)
Log(Income)	-0.423 (0.021)	-0.437 (0.022)	-0.450 (0.022)
Education	-0.446 (0.016)	-0.454 (0.017)	-0.463 (0.017)
Dinczero	-3.180 (0.190)	-3.274 (0.199)	-3.372 (0.203)

are the same. Similar results are also obtained using regression and probit models. From the estimates we see that females and old people are more risk averse, while individuals with a higher education level or a higher income are less risk averse. The estimate of the parameter for the dummy indicating that the individual has no personal income indicates that the level of risk aversion for such an individual is similar to the risk aversion of an individual with an average income. The extent to which a different value for the index results in different behavior will be discussed at the end of this section.

Table 5: Minimum distance estimates for β. Standard errors in parentheses.

	All questions	only PE questions
Female	1 (——)	1 (——)
Age	0.024 (0.001)	0.031 (0.001)
Log(Income)	-0.293 (0.007)	-0.436 (0.008)
Education	-0.321 (0.005)	-0.454 (0.006)
Dinczero	-2.160 (0.063)	-3.273 (0.070)

Since we are interested in a single measure of risk aversion, we combine the estimates for the questions using minimum distance with an optimal weighting matrix. The first column of Table 5 presents the resulting estimate for β using the choice and the probability equivalence questions jointly. However, when we test the hypothesis that the original estimates for the questions are estimates of the same β, this hypothesis is strongly rejected.[4] The correlation between the estimated indices, however, is high, ranging from 0.56 to 0.99. Looking at the questions, there is a large difference between the choice questions and the probability equivalence questions. The choice questions themselves also apply to different aspects of individual decision making, which were discussed when we presented the questions. When we test whether the estimates for β that we derived from the choice questions are the

[4]The test is based on the scaled sum of squares of the differences between the original estimates and the minimum distance estimate and follows a χ^2 distribution.

same, this hypothesis is again strongly rejected. For the probability equivalence questions we cannot reject the hypothesis that they are estimates for a unique β. The minimum distance estimate for β using only the probability equivalence questions is presented in the second column of Table 5. The hypothesis that the index for one of the choice questions was the same as the joint index for the probability equivalence questions was rejected for each of the choice questions. In the rest of this section we will refer to the index based on the probability equivalence questions as the index of risk aversion. We will denote this index with $x'\beta^{PE}$. The fact that the probability equivalence questions, in general, induce high levels of risk aversion has no consequences, since the index only represents an ordering of the respondents with respect to their level of risk aversion.

We now give a possible interpretation of the most important differences between the parameter estimates for the different questions. In general, we can say that the effect of education and income on the index is smaller in the choice questions than in the probability equivalence questions. The effect of age is similar for all the questions except for CH^4, where the effect is much stronger, even if we compare it with the other coefficients. Subproportionality of the probability weighting function might be more important for old people.

For CH^5 the estimated parameters for income and education in the index are different from the other questions. Since this question involves a loss, we can interpret the observed difference in the parameter estimate as being related to an index for loss aversion. Loss aversion means that losses have a larger disutility than gains of the same magnitude. For a detailed description of loss aversion and its possible causes see Kahneman et al. (1991). The difference between the estimate for CH^5 and the other estimates then implies that loss aversion is less decreasing with income and education level than an individual's risk attitude.

We estimated our model under the assumption that the answers to the questions depend on the level of risk aversion and that the answers are increasing in risk

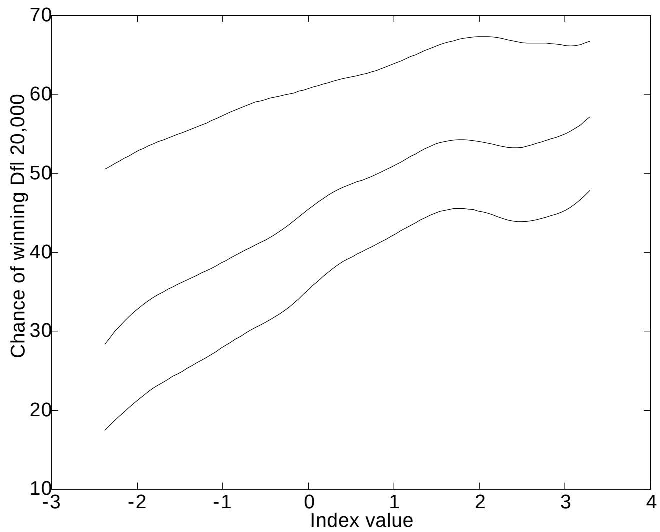

Figure 1: Estimated conditional expectation of $PE^{1,}$ PE^2, and PE^3 as a function of the index.

aversion. In Figure 1 a plot is made of $E\{PE^q|x'\beta^{PE}\}$ for the three questions, where, as the consistency requirement indicates, the lowest line is for PE^1, the middle line for PE^2, and the highest line for PE^3. We do not include confidence bands in the figure, since they make the figure unreadable, but based on uniform confidence bands we can conclude that the monotonicity of neither of the three lines is rejected and that the three conditional expectations are significantly different from each other. An estimate of the density of the index value of the respondents in our sample is presented in Figure 2. Given the fact that the density of the index is well spread over the interval $[-2, 3]$,we can conclude from Figure 1 that there is substantial variation in the individual's risk attitudes in our sample. There is, however, also a lot of unexplained variation. A measure of fit[5] based on the usual R^2 measure obtains values of 0.023, 0.067, and 0.092 for PE^3, PE^2, and PE^1, respectively, indicating that we explain relatively more for the first two questions, PE^1 and PE^2.

Although we had to reject the hypothesis that we could use a single index to

[5]We define the measure of fit as $1 - \frac{V\{PE_i^q - E\{PE_i^q|x_i'\beta^{PE}\}\}}{V\{PE_i^q\}}$. Here $PE_i^q - E\{PE_i^q|x_i'\beta^{PE}\}$ is the prediction error for respondent i.

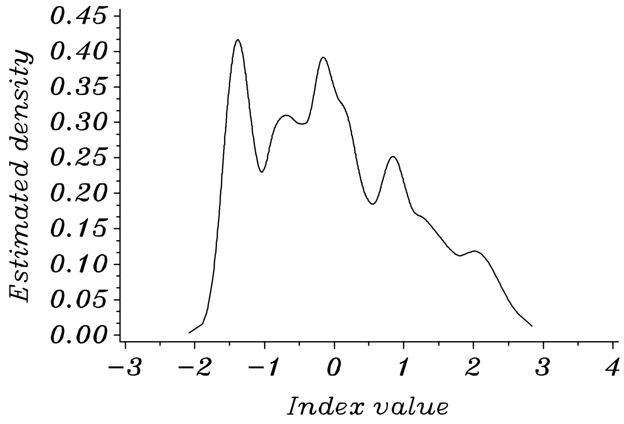

Figure 2: Estimated density for the index, $x'\beta^{PE}$.

model all the questions, it can still be the case that the measure of attitude towards risk derived from the probability equivalence questions has some predictive power for the choice questions. If this is not the case, it might not make a lot of sense to pay attention to such a measure, since it could be too dependent on the form of the questions and might have hardly anything to say about a more general attitude towards risk. However, if we do find significant relationships between the answers to the choice questions and the index based on the probability equivalence questions, we can interpret the index as a general measure of risk aversion. To check whether the index also has some predictive power for the choice questions in the sense that a higher index for an individual is related to a higher probability of choosing the safest option we performed nonparametric regressions of the answers to the choice questions on the index derived from the probability equivalence questions. The results of these regressions are presented in Figure 3.

For each question the estimated conditional expectation tends to increase. We used uniform confidence bands to formally test the null hypothesis that the functions are flat horizontal lines. For CH^1 and CH^5 we are not able to reject this hypothesis. For the other three questions we could reject this hypothesis and the conditional expectation for each of these questions significantly depends on our measure of

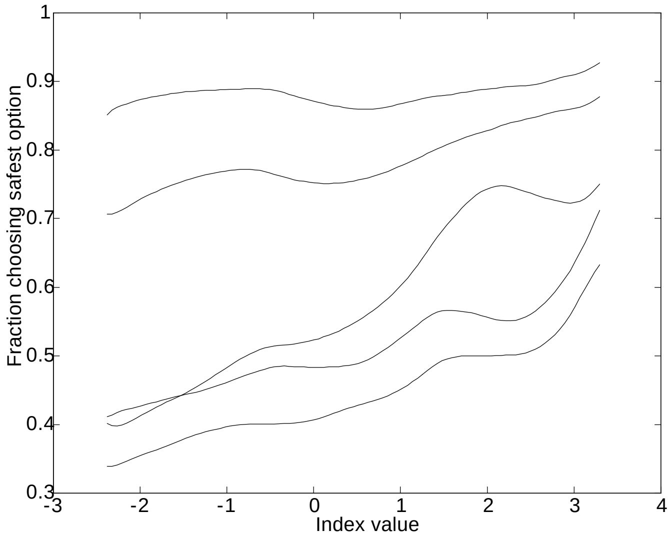

Figure 3: Estimated conditional expectation for the choice questions. From top to bottom we have CH^5, CH^1, CH^2, CH^3, and CH^4.

risk attitude. We can thus conclude that even though we could not model all the questions with a single index, we can still obtain an index that is related to all the questions and that can thus be interpreted as a general measure of risk aversion.

However, from the relationships that are depicted in Figures 1 and 3 we cannot conclude very much about the way the underlying decision process changes if the value of the index changes. The next section presents a structural model of the individual's decision making process, which will help us to interpret the results discussed above.

5 A structural model for the individual's decision making process

In economics the basic tool to deal with decision making under uncertainty is the expected utility model. This means that preferences over probability distributions can be represented by an expected utility function $E\{u(x)\} \equiv \int_X u(x)\mathrm{d}F(x)$, where $u(x)$ is a utility function and the expectation is taken with respect to the probability distribution, F. Two well known measures of risk aversion are derived by Arrow (1965) and Pratt (1964). Let $u(x)$ be the utility function of an individual, then $-\frac{xu''(x)}{u'(x)}$ is a (local) measure of relative risk aversion, while $-\frac{u''(x)}{u'(x)}$ is a (local) measure of absolute risk aversion. These are the types of concepts we are interested in. However, a lot of systematic violations of expected utility maximizing behavior have been found using questions on lotteries, one of the most famous being the Allais paradox (Allais, 1953). A good description of the evidence can be found in Kahneman, Slovic and Tversky (1982), while more recent surveys are found in Machina (1987) and Camerer (1989). With this evidence in mind various theories have been developed to explain the observed deviations from expected utility theory. Typical examples of these theories are given by Bell (1982, 1985), Gul (1995), Kahneman and Tversky (1979), Loomes and Sugden (1982), Machina (1982), Quiggin (1982), Tversky and Kahneman (1992) and Viscusi (1989).

We choose to model the individual's decision process by Cumulative Prospect Theory (Tversky and Kahneman (1992)), which is the modern version of Prospect Theory (Kahneman and Tversky (1979)). We prefer Cumulative Prospect Theory (CPT) over the other theories, mainly because CPT remains closest to Expected Utility Theory in the sense that the value of a certain lottery does not depend on the other lottery that is offered. Another advantage is that more general problems (for example, choices out of sets of 3 lotteries) can still be handled with CPT, while the generalizations of the other theories are not clear. Machina's (1982) theory seems rather difficult to use in an empirical application and Gul's (1995) Disappointment Aversion is, given our data, observationally equivalent with Prospect Theory for a specific form of the probability transformations.[6]

CPT provides us with a representation of preferences, defined over lotteries on a real interval. Our discussion will concentrate on prospects, which are lotteries with a finite number of possible outcomes. General prospects are denoted by P and

[6]Gul makes this observation when he discusses choices between binary lotteries on p. 677.

represent a set of n ordered outcomes $x_1 \leq \ldots \leq x_k \leq 0 \leq x_{k+1} \leq \ldots \leq x_n$, with corresponding probabilities $p_1, \ldots, p_n$.

In CPT the decision process consists of two phases: the editing phase and the evaluation phase. When deciding on the choice between lotteries an individual starts with the editing phase. The major operations in this phase are coding, combination, and cancellation. In this phase the decision problem is also simplified. Dominated lotteries are rejected, very small probability events deleted and probabilities and outcomes rounded off. This phase already explains some of the expected utility anomalies reported in the literature. Even though we have very simple lotteries, this phase might be relevant, since there might be shifts in reference points or other types of framing effects. Evidence on the presence of framing effects is documented in, for example, Hershey and Schoemaker (1985) for the difference between probability equivalence questions and certainty equivalence questions. Recently, Seidl and Traub (1997) discuss the differences and possible causes for a broader range of questions. A more general discussion about framing effects can be found in, for example, Kahneman, Slovic, and Tversky (1982) and Tversky and Kahneman (1991).

In the evaluation phase CPT preferences over (edited) prospects are represented by a sign and rank-dependent functional, $V(P)$, which is defined as follows:

$$\begin{aligned} V(P) = {} & \Sigma_{j=1}^{k} \left(w^{-}(\Sigma_{i=1}^{j} p_i) - w^{-}(\Sigma_{i=1}^{j-1} p_i) \right) v(x_j) \\ & + \Sigma_{j=k+1}^{n} \left(w^{+}(\Sigma_{i=j}^{n} p_i) - w^{+}(\Sigma_{i=j+1}^{n} p_i) \right) v(x_j) \qquad (1) \end{aligned}$$

Here $v(x)$ represents a *value function* for money outcomes, which is strictly increasing and continuous. $v(0)$ is set to 0 as a normalization. $w^{+}(.) : [0,1] \to [0,1]$ and $w^{-}(.) : [0,1] \to [0,1]$ are *probability weighting functions,* which transform the cumulative distribution function to a new function, similar to a distribution function. $w^{+}(.)$ is used for outcomes in the positive domain, while $w^{-}(.)$ is used for negative outcomes. Both $w^{+}(.)$ and $w^{-}(.)$ are strictly increasing and $w^{+}(0) = w^{-}(0) = 0$ and $w^{+}(1) = w^{-}(1) = 1$.

The weights assigned to the values of the outcomes when evaluating a lottery are called decision weights. The decision weights result from the transformed cumulative distribution function in the same way as probabilities result from the cumulative distribution function. For example, for a positive outcome j the decision weight equals $\left(w^+(\Sigma_{i=j}^n p_i) - w^+(\Sigma_{i=j+1}^n p_i)\right)$.

One of the important features of CPT is that $v(.)$ is defined over the lottery prizes, which are changes in wealth and not final wealth. The model uses a reference point and thus allows the magnitude of the effect of a gain to be different from the effect of an equally large loss. Individuals now are supposed to choose the lottery with the highest V value.

There has already been extensive research (see, among others, Tversky and Kahneman (1992) and Tversky and Fox (1995)) on the properties of the decision weights as transformed probabilities and, in general, it is found that small probabilities are overweighted while larger probabilities are underweighted. An example of a probability weighting function, $w^+(p)$ or $w^-(p)$, is given in Figure 4.

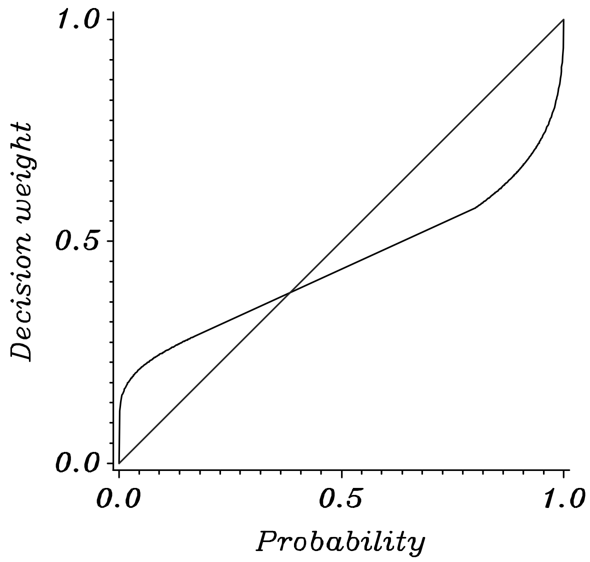

Figure 4: An example of a probability weighting function

Behavior towards risk is in CPT, unlike in expected utility theory, determined not only by the value function, $v(.)$, but also by the transformation of the probabilities,

$w(.)$. There is some debate on what defines risk attitude within non-expected utility models, but for CPT a clear discussion of the two aspects is given by Wakker (1994). He separates the effects of risk aversion in terms of $v(.)$ and $w(.)$, where the effect of $v(.)$ is called decreasing, constant or increasing marginal utility, while the effect of $w(.)$ is called probabilistic risk aversion. The effect of $v(.)$ can be characterized by the usual Arrow-Pratt measure $-\frac{v''(x)}{v'(x)}$, while the effect of $w(.)$ is measured by its convexity which can be expressed similarly. Both a stronger decrease of marginal utility[7] and a more convex transformation from probabilities to decision weights cause an individual to be more averse towards risk. The total effect depends on the prospect under consideration.

6 An empirical model for Cumulative Prospect Theory

The estimation results from the reduced form model in Section 4 show that individual characteristics influence an individual's choices in the questions that are asked, but it does not provide us with full information about the way this happens. Possibly an individual's value function varies with this index, but also the way probabilities are transformed into decision weights can be different across individuals. From the semiparametric estimation results we concluded that a single index may be too restrictive to model all the questions adequately. To test whether a model using different indices for the value and probability transformation function is able to fit all questions, one would like to use a structural model with separate indices for each of the questions. With such a model one can test whether the indices are the same for the different questions. Unfortunately, however, we cannot identify the decision weights, the value function and framing effects separately on the basis of one choice.

[7]A stronger decrease of marginal utility for individual 2 compared to individual 1 is equivalent with $v_2 = \phi \circ v_1$, with ϕ a continuous, concave and strictly increasing function. We will call v_2 more concave than v_1.

For this reason we will not use the choice questions in the analysis that follows. For the probability equivalence questions the semiparametric estimation results showed that we can use the same index for the three questions. We use these three questions to determine the way in which the observed characteristics influence the decisions an individual makes. We use Cumulative Prospect Theory (CPT) to model the individual's decision process.

The most general specification of the CPT preference representation (1) for prospects with one positive outcome, x, with probability p and 0 otherwise is:

$$V_i(0,(1-p);x,p) \equiv w_i^+(p)v_i(x) \tag{2}$$

The subscript i indicates that the function depends on the individual. Since both w and v might vary across individuals we want to allow both functions to depend on an individual's observed characteristics. We thus allow each function to depend separately on an index, $x_i'\beta_v$ for v and $x_i'\beta_w$ for w, where x_i is a vector of observed characteristics and β_v and β_w are vectors of parameters that have to be estimated. Linearity of the index is not such a strong assumption since the index is allowed to enter the model nonlinearly. For both indices we set the parameter for gender to 1 as a normalization. If β_v and β_w are the same, the model is a single index model as in Section 3.

For the choice of the functional form of the value function, v_i in (2), we follow the approach by Tversky and Kahneman (1992), and use the power function $v_i(x) = (x)^{\alpha_i}$, where we allow α_i to depend on the index $x_i'\beta_v$ quadratically,[8] so $\alpha_i = \alpha^0 + \alpha^1(x_i'\beta_v) + \alpha^2(x_i'\beta_v)^2$.

For the probability weighting function, w_i^+ in (2), we take the specification that is implied by the axiomatization of Prelec (1998), Proposition 1(A), so $w_i^+(p) = \exp(-(-\ln p)^{\gamma_i})$, where we allow γ_i to depend on the index for the probability weighting function $x_i'\beta_w$ in an affine way, so $\gamma_i = \gamma^0 + \gamma^1(x_i'\beta_w)$. The more gen-

[8] The indices are calculated using centered explanatory variables so the average for the indices is 0.

eral form of $w^+(p)$ as presented by Prelec in proposition 1(B) is not identified given our choice for $v(x)$.

In general, the effect of γ_i in the probability weighting function on an individual's risk attitude is not straightforward, since it is not directly linked to the convexity of the probability weighting function. The effect of α_i in the value function, however, is clear. A lower value of α_i implies a more concave value function and thus more aversion towards risk. We define more risk aversion as having a more concave value function and thus a lower value of α_i.

For the probability equivalence questions we assume that the respondents answered the questions in such a way that they are indifferent between the amount of money for sure and a lottery with a prize of Dfl 20,000, which might be won with the probability they answer. This implies that, for example, PE^1 satisfies the following equality: $w^+(\frac{PE^1}{100})v(20,000) = v(200)$. However, given the empirical evidence on framing effects, we want to allow for such effects. We cannot distinguish between the framing effects of the type of question we use, compared to other types of questions such as certainty equivalence questions, but we can identify differences between the questions. The framing effects we will estimate are based on the differences between the questions that are not explained by the CPT model. The estimates of the CPT parameters, especially the level of risk aversion, are still influenced by the fact that we use probability equivalence questions instead of an other type of question.

The estimated framing effects might contain systematic differences due to misspecification of our model, but in general framing effects are the result of a different interpretation by the respondents due to different questions. With our questions the respondents might adjust their reference point, since there is the possibility of having an amount of money for sure. If this is the case, this causes systematic differences between the model's predictions and actual behavior. The extent to which the reference point is adjusted might even depend on the amount of money. Such behavior is difficult to model explicitly and we allow for such factors by allowing the level of α_i to vary over the questions. We assume that the framing effects, denoted by f^1, f^2, and f^3 for PE^1, PE^2, and PE^3 respectively, are additive constants to α_i in each question. For PE^1, α_i increases with f^1 and similarly f^2 and f^3 are added to α_i for PE^2 and PE^3, respectively. Since the framing effects are not identified separately from α^0, we assume that the average framing effect equals zero. We thus set $f^1 + f^2 + f^3 = 0$ as an identifying restriction. We will distinguish between v_i, which is the individual's value function, and v_i^f, which is the individual's value function taking framing effects into account. v_i is the same for each question, while v_i^f can vary across the questions due to the framing effects. Notice that we assume that the probability weighting function is the same for each question and not affected by framing effects. This is only by assumption and the same results can be obtained if we fix the value function and allow the probability weighting function to vary across the questions in a specific manner. This should be taken into account when interpreting the results. The estimation results for the indices that control the variation between individuals are not influenced by the assumption that only the value function is influence by the framing effects.

To allow for measurement error and unobserved heterogeneity we introduce a random component with a lognormal distribution in our model . For PE^1 our final model including both the random component and the framing effects will be:

$$w^+(\frac{PE^1}{100})v^f(20,000)\eta_1 = v^f(200), \tag{3}$$

with $\eta_1|x \sim \text{Lognormal}(0,\sigma_1^2) = LN(0,\sigma_1^2)$

$$w^+(p) = \exp\{-(-\ln(p))^{\gamma_0+\gamma_1(x'\beta_w)}\}$$

$$v^f(x) = x^{\alpha_0+\alpha_1(x'\beta_v)+\alpha_2(x'\beta_v)^2+f^1}$$

The same specification is used for the other two questions with the framing effect and the value 200 replaced by the corresponding values for the other questions. To take into account the fact that we observe three questions for each individual and to allow for unobserved heterogeneity and measurement error in our measure of risk aversion we specify a general correlation structure between the errors for the different questions. The distribution of $\eta = (\eta_1, \eta_2, \eta_3)'$ is $LN(0, \Sigma)$, with Σ a full covariance matrix and η is assumed to be independent from x.

We estimated the model using maximum likelihood. Table 6 presents the es-

Table 6: Estimation results for β_v and β_π. Standard errors in parentheses.

	β_v	β_w
Female	1	1
Age	0.035 (0.009)	0.075 (0.026)
Log(inc)	-0.392 (0.114)	-0.581 (0.272)
Edu	-0.453 (0.096)	0.023 (0.081)
Dinczero	-2.782 (1.054)	-4.731 (2.367)

timates for β_v and β_w. The estimates for β_v are very similar to the estimates we obtained with the semiparametric estimation techniques. The estimate for β_w shows that the observed individual characteristics have a different effect on the probability weighting function. The effects of age and income are larger relative to the effect of gender, while the effect of education is insignificant. The estimates indicate that the difference between the true probabilities and the decision weights that are used in the decision making process is larger for females and older people. Individuals with a higher income transform the probabilities to a lesser degree. The fact that the probability weighting function and the value function depend on the individual's characteristics through different indices conflicts with the semiparametric model, which is based on the assumption that there is only one index influencing the individual's decision process. The fact that the index for the value function, $x_i'\beta_v$, is so similar to the semiparametrically estimated index, could reflect the small effect of β_w. The explained variation due to variation in $x_i'\beta_w$ is very small. If we take as a measure of fit the variance of the point forecasts[9] relative to the variance of the answers, this measure is less than 0.002 for each question if we set $x_i'\beta_v = 0$. If we set

[9]The point forecast for PE^1, for example, is $\pi^{-1}(\frac{v(200)}{v(20{,}000)})$.

$x_i'\beta_w = 0$, this measure is 0.02, 0.07, and 0.12 for PE^3, PE^2, and PE^1, respectively, so the variation due to $x_i'\beta_v$ is much larger than the variation due to $x_i'\beta_w$. The results from the semiparametric model are thus very closely related to the variation in v, giving us a possible interpretation for the semiparametrically estimated index $x_i'\beta^{PE}$. Still the restriction on the model that the probability weighting function is the same for each individual is strongly rejected.

Table 7: Estimation results for the CPT parameters. Standard errors in parentheses.

Parameters for w		Parameters for v		
γ^0	γ^1	α^0	α^1	α^2
0.393 (0.006)	-0.016 (0.005)	0.357 (0.003)	-0.021 (0.004)	0.003 (0.001)

The parameters determining the shape of the probability weighting function and the value function are presented in Table 7. For the decision weighting function we see that the estimate for γ^0, the parameter that determines the level of γ, is 0.393 and significantly different from 1, which is the case where the decision weights would be equal to the probabilities and expected utility would be valid. The estimate of γ^1, the parameter relating to the variation of γ across individuals, is significant and results in a variation of γ_i between 0.32 and 0.44. Given the estimated value function expected utility is strongly rejected. With these estimates we cannot conclude anything about the level of probabilistic risk aversion as it is defined by Wakker (1994), since γ_i is not related to the convexity of the probability transformation.

The estimates for α imply that, with the index $x_i'\beta_v$ varying between -2.7 and 3.7, for each individual $\alpha_i < 1$, which indicates that individuals have decreasing marginal utility. The estimates for α^1 and α^2 show that there is significant variation in the level of α_i. The size of the variation is, however, difficult to derive from these numbers. To give an idea about the variation across individuals and the predictions of our model, we plotted v_i for the three questions (with corresponding amounts of money) for different values of the index, $x_i'\beta_v$. This is done in the left panel of Figure 5. We normalized the scale of v_i such that $v_i(20,000) = 1$. With this normalization

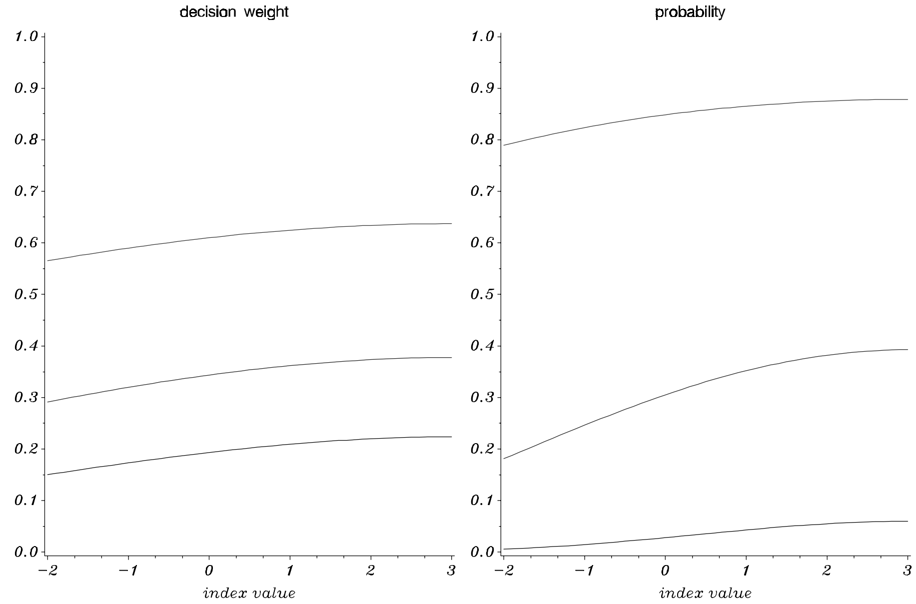

Figure 5: Predictions from the estimated CPT model excluding framing effects.

v_i, evaluated at the amount of money that is relevant for the question, equals the decision weight that is needed to be indifferent between the lottery ticket and that amount of money for sure as follows from (3). The variation in the predicted values for the decision weights in this figure is the effect of differences across individuals in the value function.

What we actually observe in the data are not the decision weights, but probabilities. In the right hand panel we plotted the probabilities that correspond to the decision weights in the left hand panel.[10] Due to the transformation of the probabilities to decision weights there is more variation in the answers than would have been the case is respondents did not transform the probabilities.

Table 8: Estimation results for the framing effects. Standard errors in parentheses.

f^1	f^2	f^3
-0.125 (0.002)	-0.048 (0.001)	0.174 (0.003)

The estimates for the framing effects of the differences between the questions are

[10] Since the probability weighting function differed across individuals through a second index, $x_i'\beta_\pi$, we set this index to 0 and used the 'average' probability weighting function with $\gamma = 0.393$.

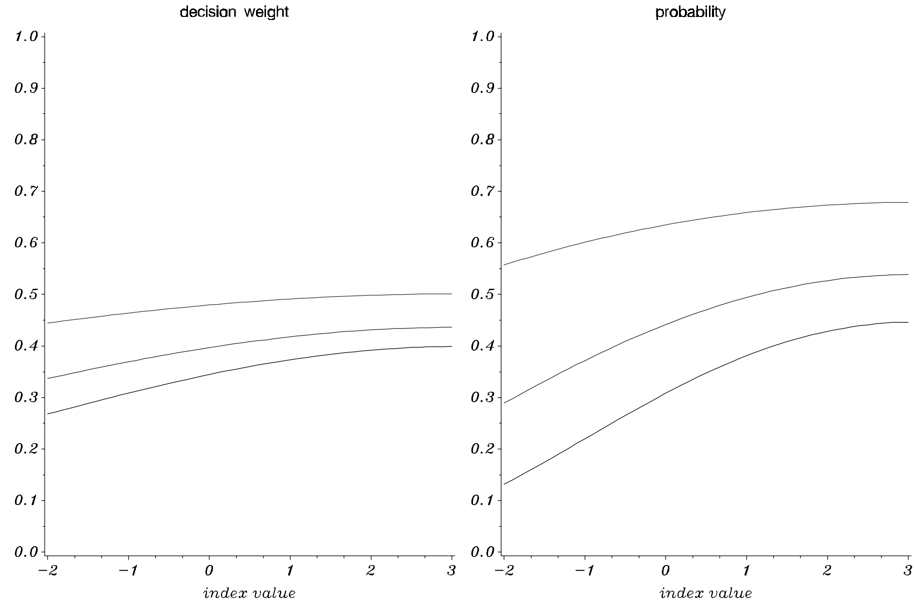

Figure 6: Predictions from the estimated CPT model including framing effects.

presented in Table 8. The framing effects are highly significant and imply higher answers for PE^1 and PE^2, while for PE^3 the answers are lower than without the framing effects. One of the reasons for this could be that our choice of functional forms is wrong, but a shift in reference points induced by the different amounts in the questions seems a better explanation. Making the reference point endogenous in the model, however, is rather difficult.

Figure 6 presents the same model predictions as Figure 5, but now v_i^f is used instead of v_i. The value function that is used to evaluate the different lotteries thus depends on the question. Comparing Figure 5 with Figure 6 we see that there is a clear need to understand framing effects in more detail, since they have a large impact on model predictions. We can conclude from these figures, however, that there is substantial variation in attitude towards risk across individuals and that we can predict part of this variation.

The estimates for the parameters in Σ are presented in Table 9. The correlation between the errors for the different questions for each individual are high, as could be expected. This indicates that there is still a lot of systematic variation at the individual level after we have taken out the systematic variation due to the observed

Table 9: Estimation results for the parameters in Σ.

Parameter	Estimate	Parameter	Estimate
ρ_{12}	0.831 (0.005)	σ_1	0.371 (0.009)
ρ_{23}	0.793 (0.007)	σ_2	0.363 (0.008)
ρ_{13}	0.593 (0.013)	σ_3	0.365 (0.007)

characteristics.

7 Conclusion

In this paper we investigated whether attitude towards risk is related to some commonly observed individual characteristics. Using semiparametric estimation techniques we find significant relationships between the answers to a set of questions on lotteries and age, gender, income, and education level. Females and older people have a more negative attitude towards risk, while income and education level are positively related to an individual's attitude towards risk. We focussed on the index that is derived from the three probability questions and, even though we rejected the hypothesis that we could use a single index for all the questions, we found positive relationships between the choices that are made in the choice questions and the index derived from the probability equivalence questions. It thus seems justified to use such an index as a general measure of risk aversion. Implementing this measure of risk aversion into a model for savings or asset holdings could be used to prove the usefulness of measuring individual risk aversion, but this is left for future research.

To obtain more insight into the way the decision process differs across the individuals, we estimated a parametric model based on CPT. The specification allowed the value function and the probability weighting function to depend on the observed characteristics through two separate indices. Also systematic deviations from the model, due to, for example, framing effects, are allowed for.

The probability weighting function varies systematically with age, income, and gender, while the value function depends on the observed characteristics similarly to the relationship we found with the semiparametric estimates. This gives a nice interpretation to the results from the semiparametric model: The semiparametrically estimated index seems to be related to the value function.

Using the decomposition of attitudes towards risk into decreasing/increasing marginal utility and probabilistic risk aversion our results indicate that individuals have decreasing marginal utility. Higher values for the estimated index imply a stronger decrease of marginal utility. Our specification does not allow us to say anything about probabilistic risk aversion, but the decision weights are significantly different from the true probabilities. For older people and females the difference is largest, while income has a negative effect on the difference.

These results are, however, based on a simple specification and complete identification of the influence of the value function and the probability transformation can only be based on a richer set of questions. One possibility to do this might be the use of a very large questionnaire as was done by Kahneman and Tversky, but it seems more fruitful in a survey to incorporate a shorter, but well designed set of questions. The ideas of Wakker and Deneffe (1996) on how to identify the utility function without specification of the decision weights might be a good starting point for this.

A Data

A.1 Questions

The first type of questions are the probability equivalence questions. In this type of questions, the probability is asked which would make the individual indifferent between a lottery ticket with probability p of winning 20,000 or a prespecified amount of money for sure.
The exact question is:

Imagine you have won Dfl amount$_k$ in a game. You can now choose between keeping that Dfl amount$_k$, or having a lottery ticket with a certain chance to win a prize of Dfl 20,000.

How high would that chance to win Dfl 20,000 have to be such that you would prefer the lottery ticket to keeping the Dfl amount$_k$ that you had already won?

I would prefer the lottery ticket if the chance to win the first prize would be at least.......... PE^k%

This question was asked three times, with amount$_k$ being Dfl 200, Dfl 1,000, and Dfl 5,000.

The second type of question is on choices between two opportunities, where preference for one or the other has to be stated.

The following information is given to the individuals.

You are probably familiar with games shown on television, where people win prizes and can choose between several options. For example, they can choose to keep a certain prize, or they can choose to take a chance to get a much bigger prize, at the risk of losing the prize all together.

The following questions present similar choices, concerning amounts of money. Some of the amounts are certain for you to have, others you can win in a lottery.

We would like to know which choice you would make. There are no right or wrong answers with these questions.

CH^1 We toss a coin once. You may choose one of the following two options:

- You receive Dfl 1,000 with either heads or tails
- With heads you receive Dfl 2,000, with tails you don't receive anything at all.

CH^2 Which of the following two options would you choose?

- You draw a lottery ticket with an 80% chance to win Dfl 45 (if you loose, you don't get anything at all)
- You win Dfl 30, no matter which ticket is drawn.

CH^3 Which of the following two options would you choose?

- You draw a lottery ticket with a 25% chance to win Dfl 100 (if you loose, you don't get anything at all)
- You draw a lottery ticket with a 20% chance to win Dfl 130 (if you loose, you don't get anything at all)

CH^4 Which of the following two options would you choose?

- You draw a lottery ticket with a 2% chance of winning Dfl 3000 (if you loose, you don't get anything at all)
- You draw a lottery ticket with a 1% chance of winning Dfl 6000 (if you loose, you don't get anything at all)

CH^5 We toss a coin once. Would you accept the following agreement? (yes/no)

- Heads, you win Dfl 1,500.
- Tails, you lose Dfl 1,000

A.2 Descriptive statistics

This appendix contains the definition and some descriptive statistics of the variables that are used as independent variables in the models that are estimated.

Table 10: Description of some variables

Variable	Description	Mean	Std. dev.
Age	Age (in years)	42.1	14.18
Female	Dummy; 1 if female	0.43	0.50
Education	Education level, 1,3,5	3.31	1.67
Log(Income)	Log(gross annual individual income)	8.70	4.27
Dinczero	Dummy; 1 if income equals zero	0.18	0.39

B Semiparametric estimation method

In this appendix we describe the method of estimation we use for the semiparametric model of section 3. We give a short description of the assumptions we make and the choices for the bandwidths in the semiparametric estimators.

The main assumption is that for each question the distribution of the answers for an individual i with characteristics x_i depends on x_i only though an index $x_i'\beta$. Let y_i be individual i's answer to the question under consideration. We then have that $f(y_i|x_i) = f(y_i|x_i'\beta)$, where $f(y_i|x_i)$ denotes the density function of y_i given x_i. Let $E\{y_i|x_i'\beta\}$ denote the expectation of y_i given $x_i'\beta$, then we can write the monotonicity assumption we make as $E\{y_i|x_i'\beta\} = G(x_i'\beta)$, with $G'(.) > 0$. We also use a normalization for the parameter relating to gender. Preliminary analysis showed that this variable had a significant influence on the answers, making it a valid parameter for the normalization. For the estimator to be consistent, there also needs to be at least one continuous variable that has a nonzero coefficient. Both age and income can satisfy this condition, but, due to the high correlation between these two variables, it can be the case that only one of these variables is significant and it is not clear a priori which one is.

With these assumptions and some regularity conditions we can use the rank estimator proposed by Cavanagh and Sherman (1998) (CS) to obtain a $\sqrt{N}-$consistent estimate for β in each question. The estimator of CS is defined as:

$$\hat{\beta}^{rc} = \arg\max_{\beta} \frac{1}{N} \Sigma_{i=1}^{N} y_i R_N(x_i'\beta), \tag{B.1}$$

with $R_N(x_i'\beta) \equiv \Sigma_{j=1}^{N} I\{x_i'\beta \geq x_j'\beta\}$, the rank if $x_i'\beta$. CS prove that the objective function is asymptotically smooth, even though the rank of $x_i'\beta$ is not a smooth function. The small sample properties of the estimator, however, are not so nice and optimization of the objective function turns out to be problematic in our case. To overcome the small sample problems of the estimator we smooth $R_N(x_i'\beta)$ as follows:

$$R^s_N(x_i'\beta) = \Sigma_{j=1}^N F(\frac{x_i'\beta - x_j'\beta}{h_N}), \tag{B.2}$$

with $F(.)$ the cumulative distribution function for the logistic distribution and h_N a smoothness parameter satisfying $h_N \to 0$ as $N \to \infty$.

The initial $\sqrt{N}-$consistent estimate is now defined as:

$$\hat{\beta}^{rcs} = \arg\max_\beta \frac{1}{N}\Sigma_{i=1}^N y_i R^s_N(x_i'\beta). \tag{B.3}$$

Optimization of the objective function is performed with a Simplex algorithm. This works well in practice. The estimate is not sensitive to the choice of h_N in the smoothed rank. For practical purposes we set $h_N = 0.1\sigma$, with σ the estimated standard deviation of $x_i'\beta$, although it might not be valid to let h_N depend on the estimated parameter.

With this initial estimate a semiparametrically efficient estimate is constructed using a one-step improvement as proposed by Delecroix, Härdle, and Hristache (1997). We define $L_n(\beta)$ as $\frac{1}{N}\Sigma_{i=1}^N \log(f(y_i|x_i'\beta))$, the likelihood function. Since we do not know $f(y_i|x_i'\beta)$ we have to estimate it. This is done using kernel estimates. We define $\hat{L}_n(\beta)$ as $\frac{1}{N}\Sigma_{i=1}^N \log(\hat{f}(y_i|x_i'\beta))$, with

$$\hat{f}(y|x'\beta) = \frac{\frac{1}{N}\Sigma_{i=1}^N \frac{1}{hy}K(\frac{y-y_i}{hy})\frac{1}{hx}K(\frac{(x-x_i)'\beta}{hx})}{\frac{1}{N}\Sigma_{i=1}^N \frac{1}{hx}K(\frac{(x-x_i)'\beta}{hx})} \tag{B.4}$$

The efficient estimate is now defined by

$$\hat{\beta} = \hat{\beta}^{rcs} - (\frac{\partial^2\hat{L}_n}{\partial\beta\partial\beta'}\left(\hat{\beta}^{rcs}\right))^{-1}\frac{\partial\hat{L}_n}{\partial\beta}\left(\hat{\beta}^{rcs}\right), \tag{B.5}$$

as long as $\frac{\partial^2\hat{L}_n}{\partial\beta\partial\beta'}\left(\hat{\beta}^{rc}\right)$ is negative definite. The gradient and Hessian need to be computed using fourth order kernels. In small samples this can be problematic since the density estimates can be negative. Instead of using the theoretically required fourth order kernels, we will use a variable bandwidth kernel density estimator (see Hall (1990) and Hall and Marron (1988)), which yields the same bias reduction, while at the same time the density estimate is guaranteed to be positive. Numerical derivatives are used to compute the gradient, while the Hessian is computed as the outer product of the gradient.

For the variable bandwidths we set $h_x = 0.0625\hat{f}(x'\hat{\beta}^{rcs})$ in de denominator, $h_x = 0.0625\hat{f}(y, x'\hat{\beta}^{rcs})$ in the numerator and $h_y = 0.125\hat{f}(y, x'\hat{\beta}^{rcs})$, where $\hat{f}(y, x'\hat{\beta}^{rcs})$ and $\hat{f}(x'\hat{\beta}^{rcs})$ are kernel estimates for the joint distribution of y_i and $x_i'\hat{\beta}^{rcs}$, and the marginal distribution of $x_i'\hat{\beta}^{rcs}$, respectively. Although Delecroix, Härdle and Hristache (1997) provide no theoretical justification for a data dependent bandwidth, as we use for the variable bandwidth kernel density estimator, we choose this approach on practical grounds. The advantages of a guaranteed positive density and a bias reduction that is the same as for fourth order kernels are large.

The values for the bandwidths are based on visual observation. Since the method described above uses undersmoothed bandwidths, we select bandwidths in the region where the density estimates are not very smooth. Within a large range of bandwidth choices the estimates did not vary very much. Standard errors for the estimates are also computed using numerical derivatives. They were more sensitive to the choice of bandwidths, but, for a reasonable range of bandwidths, they do not differ by more than 25% from the estimates we present here.

References.

Allais, M. (1953), "Le comportement de l'Homme Rationnel devant le Risque, Critique des Postulats et Axiomes de l'Ecole Americaine", Econometrica, 21, 503-546

Abdeallaoui, M. (1998), "Eliciting Utilities and Decision Weights under Cumulative Prospect Theory", unpublished manuscript, ENS de Cachan, France

Arrow, K. (1965), "Aspects of the Theory of Risk-bearing", Helsinki, Yrjö Jahnssonin Säätiö

Beetsma, R.M.W.J., and P.C. Schotman (1997), "Measuring Risk Attitudes in a Natural Experiment: an Empirical Analysis of the Television Game Show LINGO", CentER discussion paper no 9808, Tilburg University, Tilburg

Bell, D. (1982), "Regret in Decision Making under Uncertainty", Operations Research, 30, 961-981

Bell, D. (1985), "Putting a Premium on Regret", Management Science, 31, 117-120

Camerer, C.F. (1989), "An Experimental Test of Several Generalized Utility Theories", Journal of Risk and Uncertainty, 2, 61-104

Cavanagh, C. and R.P. Sherman (1998), "Rank Estimators for Monotonic Index Models", Journal of Econometrics, 84, 351-381

Delecroix, M., W. Härdle, and M. Hristache (1997), "Efficient Estimation in Single Index Regression", mimeo, Humboldt Universitat, Berlin

Guiso, L., T. Japelli, and D. Terlizzese (1992), "Earnings Uncertainty and Precautionary Saving", Journal of Monetary Economics, 30, 307-337

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

Hall, P. (1990), On the bias of variable bandwidth curve estimators, Biometrika, 77, 529-535

Hall, P. and Marron, J.S. (1988), Variable window width kernel estimates of probability density estimates, Probability Theory and Related Fields, 80, 37-49

Han , A.K. (1987), "Non-parametric Analysis of Generalized Regression Model", Journal of Econometrics, 35, 303-316

Härdle, W. and O. Linton (1994), "Applied Nonparametric Methods", in R. Engle and D. McFadden (eds), Handbook of Econometrics, Vol IV, North-Holland, Amsterdam, 2297-2341

Harless, D.W. and C.F. Camerer (1994), "The Predictive Utility of Generalized Expected Utility Theories", Econometrica, 62, 1251-1290

Hartog, J., N. Jonker, and H. Oosterbeek (1997), "On a Simple Measure of Risk Aversion", mimeo, University of Amsterdam, Amsterdam

Hershey, J.C. and P.J.H. Schoemaker (1985), "Probability versus Certainty Equivalence Models in Utility Measurement", Management Science, 31, 1213-1231

Horowitz, J.L. (1992), " A Smoothed Maximum Score Estimator for the Binary Response Model", Econometrica, 60, 505-531

Jullien, B., and B. Salanie (1997), "The Attitudes Towards Risk of Racetrack Bettors", mimeo, Universite des Sciences Sociales de Toulouse, Toulouse

Kahneman, D., J.L. Knetsch and R.H. Thaler (1991), "Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias", Journal of Economic Perspectives, 5, 193-206

Kahneman, D., P. Slovic and A. Tversky (eds.) (1982), Judgement under Uncertainty: Heuristics and Biases, New York, Cambridge University Press

Kahneman, D. and A. Tversky (1979), "Prospect Theory: An Analysis of Decision Under Risk", Econometrica, 47, 263-291

Klein, R.W., and R.H. Spady (1993), "An Efficient Semiparametric Estimator for Binary Response Models", Econometrica, 61, 387-421

Lattimore, P.M.J., J.R. Baker, and A.D. Witte (1992), "The Influence of Probability on Risky Choice", Journal of Economic Behavior and Organization, 17, 377-400

Lee, M-J. (1996), Methods of Moments and Semiparametric Econometrics for Limited Dependent Variable Models, Springer Verlag, New York

Loomes, G. and R. Sugden (1982), "Regret Theory: An Alternative Theory of Rational Choice Under Uncertainty", Economic Journal, 92, 805-824

Lusardi, A. (1997), "Precautionary Saving and Subjective Earnings Variance", Economics Letters, 57, 319-326

Machina, M. J. (1982), "Expected Utility Without the Independence Axiom", Econometrica, 50, 277-323

Machina, M. J. (1987), "Choice Under Uncertainty: Problems Solved and Unsolved", Journal of Economic Perspectives, 1, 121-154

Nyhus, E.K. (1996), "The VSB-CentER savings project: data collection methods, questionnaires and sampling procedures", mimeo, CentER, Tilburg

Pålsson, A.-M. (1996), "Does the Degree of Relative Risk Aversion Vary with Household Characteristics?", Journal of Economic Psychology, 17, 771-787

Pratt, J. (1964), "Risk Aversion in the Small and in the Large", Econometrica, 32, 122-136.

Prelec, D. (1998), "The Probability Weighting Function", Econometrica, 66, 497-528.

Quiggin, J. (1982), "A Theory of Anticipated Utility" Journal of Economic Behaviour and Organization, 3, 323-343

Seidl, C. and S. Traub (1997), "Assessment Biases in Utility Functions",mimeo, Christian-Albrechts-Universitat, Kiel

Tversky, A., and C.R. Fox (1995), "Weighing Risk and Uncertainty", Psychological Review, 2, 269-283

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

Viscusi, K.W. (1989), "Prospective Reference Theory: Toward an Explanation of the Paradoxes", Journal of Risk and Uncertainty, 2, 235-264

Wakker, P. (1994), "Separating Marginal Utility and Probabilistic Risk Aversion", Theory and Decision, 36, 1-44

Wakker, P., and D. Deneffe (1996), "Eliciting von Neuman-Morgenstern Utilities When Probabilities are Distorted or Unknown", Management Science, 8, 1131-1150

Warneryd, K.E. (1996), "Risk Attitudes and Risky Behavior", Journal of Economic Psychology, 17, 749-770