ETH zürich

ETH Library

Rationality on the rise

Why relative risk aversion increases with stake size

Journal Article

Author(s):
Fehr-Duda, Helga; Bruhin, Adrian; Epper, Thomas; Schubert, Renate

Publication date:
2010-04

Permanent link:
https://doi.org/10.3929/ethz-b-000016857

Rights / license:
In Copyright - Non-Commercial Use Permitted

Originally published in:
Journal of Risk and Uncertainty 40(2), https://doi.org/10.1007/s11166-010-9090-0

This page was generated automatically upon download from the ETH Zurich Research Collection.
For more information, please consult the Terms of use.

J Risk Uncertain (2010) 40:147–180
DOI 10.1007/s11166-010-9090-0

Rationality on the rise: Why relative risk aversion increases with stake size

Helga Fehr-Duda · Adrian Bruhin · Thomas Epper · Renate Schubert

Published online: 3 March 2010
© Springer Science+Business Media, LLC 2010

Abstract How does risk tolerance vary with stake size? This important question cannot be adequately answered if framing effects, nonlinear probability weighting, and heterogeneity of preference types are neglected. We show that the observed increase in relative risk aversion over gains cannot be captured by the curvature of the value function. Rather, it is predominantly driven by a change in probability weighting of a majority group of individuals who weight probabilities of high gains more conservatively. Contrary to gains, no coherent change in relative risk aversion is observed for losses. These results not only challenge expected utility theory, but also prospect theory.

Keywords Stake effects · Prospect theory · Expected utility theory · Risk aversion

JEL Classification D81 · C91

Risk is a ubiquitous feature of social and economic life. Many of our decisions, such as what trade to learn and where to live, involve risky consequences of great importance. Often these choices entail substantial monetary costs and rewards. Therefore, risk taking behavior under high stakes is a relevant area of economic research. The effect of stake size on risk tolerance has been debated since the early days of expected utility theory as economic theory is agnostic about the existence, direction and size of stake effects. In a classical paper,

H. Fehr-Duda (✉) · A. Bruhin · T. Epper · R. Schubert
Chair of Economics, ETH Zurich,
Weinbergstrasse 35, 8092 Zurich, Switzerland
e-mail: fehr@econ.gess.ethz.ch

Markowitz (1952) surmised that risk preferences are likely to reverse from risk seeking over very small stakes to risk aversion over high stakes.

Markowitz did not test his conjecture experimentally, but abundant evidence has accumulated by now showing that relative risk aversion indeed rises with stake size. However, due to the limits of experimental budgets, the bulk of the studies addressing stake sensitivity are either based on quite limited payoff ranges or, when substantial payoffs are involved, on hypothetical choices (Hogarth and Einhorn 1990; Bosch-Domenech and Silvestre 1999; Kuehberger et al. 1999; Weber and Chapman 2005; Astebro et al. 2009). While experiments with payoffs in the range of a few dollars may not reveal the effective degree of risk tolerance over truly high stakes, results based on hypothetical choices may not be informative, either, as the literature on incentive effects suggests. A large body of experimental evidence shows that it makes a difference whether subjects respond to decision situations with hypothetical or real monetary consequences: In general, subjects tend to be relatively more risk averse when real money is at stake (Smith and Walker 1993; Wilcox 1993; Beattie and Loomes 1997; Camerer and Hogarth 1999). A striking example is provided by Holt and Laury (2002, 2005) who find that, contrary to their results on real payoffs, subjects' risk aversion exhibited over hypothetical lotteries does not change with increasing stake size at all. Therefore, in order to be able to address the issue of stake sensitivity in a satisfactory way, experiments with real substantial payoffs are needed.

Not surprisingly, there are only a handful of experiments with substantial monetary incentives, typically conducted in developing countries (Grisley and Kellog 1987; Wik et al. 2004). Two prominent studies in this category, both of which investigate behavior over risky gains only, are Binswanger (1981) and Kachelmeier and Shehata (1992). Binswanger reports data on subjects from rural India with stakes amounting to a month's average income. He reaches the same conclusion as Kachelmeier and Shehata who paid up to three month's wages in Beijing, China: Relative risk aversion over gains increases significantly when stakes are raised from low payoffs to substantial ones. In agreement with these experimental findings field data on behavior in game shows, where prizes up to a million dollars can be won, also show that contestants tend to behave more conservatively when faced with higher stakes (Bombardini and Trebbi 2005; Andersen et al. 2006b; Baltussen et al. 2008; Post et al. 2008).[1] Therefore, the empirical evidence so far seems to confirm Markowitz's conjecture of increasing relative risk aversion.

[1] Arguably, game shows provide a decision environment which differs radically from everyday situations. Contestants face a once-or-never opportunity to win an extremely high amount of money, and they have to take their decisions under time pressure and under the scrutiny of a large audience. Estimates of overall risk aversion in game shows suggest a relatively high level of risk tolerance which is most likely not representative of risk attitudes in "normal" situations. Nevertheless, in line with the experimental evidence, when facing comparatively higher stakes contestants become relatively more risk averse.

In decision situations with clearly defined monetary outcomes and objectively given probabilities, such as in controlled experimental settings with context, framing and response mode held constant, changes in risk taking behavior should only result from a change in the evaluation of outcomes and/or probabilities. Most economists would attribute this change in relative risk aversion to the characteristics of the utility for money, and would search for suitable functional forms that are able to accommodate this behavioral pattern. Little is known about the underlying forces of the increase in relative risk aversion, however. In particular, it is not clear whether the change in risk tolerance is actually a consequence of the way people value low versus high amounts of money or whether probability weighting is sensitive to stake size.

Most previous studies focused on the overall effect of stake size on risk taking, ignoring probability weighting. In their seminal contribution, Tversky and Kahneman (1992) already suspected that probability weights might be responsive to the level of outcomes, but they questioned whether increasing the complexity of decision theory was worth the costs of such an endeavor. Some preliminary findings indeed suggest that probability weights may be stake dependent (Camerer 1991). One of the few experimental studies that explicitly addressed this possibility is Kachelmeier and Shehata (1992). They find that there is an interaction effect of stake size with probability level, namely that stake-sensitivity is greater for smaller probabilities, but their data set is not sufficiently rich to draw any conclusions on the relative contributions of outcome valuation and probability weighting to the change in risk attitudes. Etchart-Vincent (2004), on the other hand, who directly investigates the stake dependence of probability weights under hypothetical losses does not detect any clear stake effect, which may be due to hypothetical bias discussed above. To summarize, relative risk aversion over real gains increases significantly with stake size. Whether utility for money or probability weighting is the driving force behind this change has not been systematically studied so far. Moreover, evidence on losses is scarce and not conclusive.

In order to close this gap, we analyze comprehensive choice data stemming from an experiment conducted in Beijing in 2005. The experimental subjects had to take decisions over substantial real monetary stakes with maximum payoffs amounting to more than an average subject's monthly income. In total, subjects were presented with 28 lotteries over gains and another 28 otherwise identical lotteries framed as losses in order to be able to investigate the effect of increasing stake size on relative risk aversion in both decision domains. To disentangle the effects of stake size on the valuation of monetary outcomes and probability weighting, we estimated the parameters of a flexible sign- and rank-dependent decision model, which nests expected utility theory as a special case. Furthermore, as average estimates may gloss over potentially important differences in individual behavior we account for the existence of heterogeneous preference types. We estimated a finite mixture regression model, which assigns each individual to one of several distinct behavioral types and provides type-specific parameter estimates for the underlying decision model (El-Gamal and Grether 1995; Stahl and Wilson 1995; Houser et al.

2004). Its ability to parsimoniously characterize distinct behavioral types is probably the most attractive feature of finite mixture models.

The following results emerge from our analysis. First, at the level of observed behavior, we find a strong and significant difference between subjects' evaluations of risky gains and losses. Whereas observed certainty equivalents over gains exhibit significantly increasing relative risk aversion, there is no coherent stake-dependent pattern in subjects' behavior over the same lotteries presented as losses, implying a significant framing effect.

Second, model estimates show that, contrary to many economists' expectations, value function parameters remain stable over increased stakes in both decision domains, implying that the observed increase in average relative risk aversion over gains cannot be explained by changing attitudes towards monetary outcomes. Rather, it can be predominantly attributed to a change in probability weighting. The probability weighting function for high gains is characterized by substantially lower probability weights over a wide range of probabilities than the respective function for low stakes, entailing less optimistic lottery evaluation and, thus, greater relative risk aversion at high stakes. This change is particularly pronounced for smaller probabilities for which the average decision maker tends to be risk seeking, corroborating the findings by Kachelmeier and Shehata (1992). In the loss domain, however, no such change in probability weights can be inferred from the estimates, consistent with the observed pattern of behavior.

Third, a model allowing for heterogeneity of preference types is clearly preferable to a representative-agent model. We find two distinct behavioral groups: The majority of about 73% of the subjects exhibit an inverted S-shaped probability weighting curve, whereas the minority can essentially be characterized as expected value maximizers. Furthermore, we show that the observed increase in average relative risk aversion over gains can exclusively be attributed to a change in behavior by the majority group of decision makers, who evaluate high-stake prospects more cautiously by putting lower weights on stated gain probabilities. In contrast, the minority type's behavior is not affected by rising stakes at all.

Our results entail material consequences for decision theory as well as applied economics. The first two findings, the framing effect as well as the probability weighting function as carrier of changing risk attitudes, effectively rule out expected utility theory as a candidate for explaining increasing relative risk aversion. Since it is the probability weights that are responsible for the change in relative risk aversion, more flexible utility functions cannot adequately solve the problem of modeling increasing risk aversion. While the observed differences between the evaluation of gains and losses, in principle, lends support to sign-dependent decision models, such as prospect theory, stake dependence of probability weights, however, calls theories based on stake-invariant probability weights into question.

The third finding poses a challenge to type-independent models of choice under risk, which might be prone to aggregation bias. We show that the vast heterogeneity in individual risk taking behavior, typically found in choice data

(Hey and Orme 1994; Gonzalez and Wu 1999; Cohen and Einav 2007), is substantive in the sense that a single preference model is unable to adequately describe behavior. If maximization of expected utility is accepted as normative standard of rational behavior the majority of the individuals clearly deviate from rationality. The heterogeneity of preference types may render policy recommendations based on average parameter estimates inappropriate when strategic interactions among market participants play a role. As the literature on individual irrationality and aggregate outcomes has shown (Haltiwanger and Waldman 1985, 1989; Fehr and Tyran 2005, 2008; Camerer and Fehr 2006), the mix of behavioral types may be crucial for the market outcome and, depending on the nature of the strategic interdependence, the behavior of even a minority of players may be decisive for the aggregate outcome. Furthermore, regulatory policy should be designed in such a way that it creates large benefits for those who make errors, while imposing little or no harm on those who are fully rational (Camerer et al. 2003). Obviously, total net benefits of a regulatory policy measure depend not only on the costs for the rational citizens and the benefits for the irrational ones, but also on the relative numbers of rational and irrational types in the population. Therefore, knowing the mix of behavioral types is important for designing cost efficient programs and regulations.

To the best of our knowledge, this is the first study that provides a systematic examination of stake effects on probability weights for real substantial payoffs. Neither are we aware of any other study that examines the relevance of framing and type heterogeneity for the impact of stakes on risk tolerance.

The remainder of the paper is structured as follows. Section 1 describes the experimental design and procedures. The decision model applied to the experimental data as well as the finite mixture regression model are presented in Section 2. The results of the estimation procedure are discussed in Section 3. Section 4 concludes the paper.

1 Experiment

In the following section, the experimental setup and procedures are described. The experiment took place in Beijing in November 2005. The subjects were recruited by flier distributed at the campuses of Peking University and Tsinhua University. Interested people had to register by email for one of two sessions conducted on the same day. Subjects were selected to guarantee a balanced distribution of genders and fields of study. In total, 153 subjects' responses were analyzed.

The experiment served to elicit certainty equivalents for 56 two-outcome lotteries over a wide range of outcomes and probabilities.[2] Twenty-eight lotteries offered low-stake outcomes between 4 and 55 Chinese Yuan (CHN)

[2]Instructions are available upon request.

Table 1 Gain lotteries (x_1, p; x_2)

p	x_1	x_2	p	x_1	x_2	p	x_1	x_2
0.05	15	4	0.25	250	65	0.75	250	65
0.05	20	7	0.25	320	130	0.75	320	130
0.05	55	20	0.50	7	4	0.90	7	4
0.05	250	65	0.50	15	4	0.90	130	65
0.05	320	130	0.50	20	7	0.95	15	4
0.05	950	320	0.50	130	65	0.95	20	7
0.10	7	4	0.50	250	65	0.95	250	65
0.10	130	65	0.50	320	130	0.95	320	130
0.25	15	4	0.75	15	4			
0.25	20	7	0.75	20	7			

Outcomes x_1 and x_2 are denominated in Chinese Yuan. p denotes the probability of the higher gain

with an expected payoff of about 16 CHN, approximately equal to the going hourly wage rate. Another 28 lotteries entailed high-stake outcomes from 65 to 950 CHN. Overall, payoffs spanned the range of 0.25 to approximately 60 hourly wages.[3] The high-stake lotteries were constructed from the low-stake ones by inflating the outcomes by approximately the same factor such that reasonable integer numbers were obtained and the highest obtainable payoff amounted to a student's monthly allowance, amounting to less than 1,000 CHN for the majority of students. Average total earnings per subject summed to approximately 323 CHN, including a show up fee of 20 CHN. Monetary incentives were substantial given the subjects' average monthly disposable income of about 700 CHN.

Probabilities of the lotteries' higher gain or loss varied from 5% to 95%. One half of the lotteries were framed as choices between risky and certain gains ("gain domain"); the same 28 decisions were also presented as choices between risky and certain losses ("loss domain"). For each lottery in the loss domain, subjects were provided with a specific endowment which served to cover their potential losses. These initial endowments rendered the expected payoff for each loss lottery equal to the expected payoff of an equivalent gain lottery. The set of gain lotteries is presented in Table 1.

Subjects were entitled to one random draw from their low-stake decisions and one random draw from their high-stake decisions. In order to preclude order effects, low-stake and high-stake lotteries were intermixed and appeared in random order in a booklet containing the decision sheets.

For each lottery, a decision sheet, such as presented in Fig. 1, contained the specifics of the lottery and a list of 20 equally spaced certain outcomes ranging from the lottery's maximum payoff to the lottery's minimum payoff. Subjects had to indicate whether they preferred the lottery or the certain payoff for each row of the decision sheet. The lottery's certainty equivalent was then

[3] At the time of the experiment one Chinese Yuan equaled about 0.12 U.S. Dollars (USD). Low-stake outcomes ranged from 0.48 to 6.6 USD, high-stake outcomes from USD 7.8 to 114 USD.

Decision situation: 22							
	Option A	Your Choice:					Option B Guaranteed payoff amounting to:
1	Profit of CHN 20 with probability 25% and profit of CHN 0 with probability 75%	A			o	B	20
2		A			o	B	19
3		A			o	B	18
4		A			o	B	17
5		A			o	B	16
6		A			o	B	15
7		A			o	B	14
8		A	o			B	13
9		A	o			B	12
10		A	o			B	11
11		A	o			B	10
12		A	o			B	9
13		A	o			B	8
14		A	o			B	7
15		A	o			B	6
16		A	o			B	5
17		A	o			B	4
18		A	o			B	3
19		A	o			B	2
20		A	o			B	1

Fig. 1 Design of the decision sheet

calculated as the arithmetic mean of the smallest certain amount preferred to the lottery and the subsequent certain amount on the list. For example, if a subject's preferences corresponded to the small circles in Fig. 1, her certainty equivalent would amount to 13.5 CHN. This elicitation procedure has been widely used in the experimental literature (Tversky and Kahneman 1992). We chose this method because it is transparent, easy to understand and well suited for a paper-and-pencil experiment.[4]

Before subjects were permitted to start working on the experimental decisions, they were presented with two hypothetical choices to become familiar with the procedure. Subjects could work at their own speed. The vast majority of them needed considerably less than 90 min to complete the experiment. After completion of the experimental tasks and a complementary socioeconomic questionnaire one of each subject's low-stake as well as high-stake choices were randomly selected for payment. Subjects executed the random draws themselves and were paid in private afterwards.

2 Econometric model

This section discusses the specification of the econometric model, which is based on several building blocks: first, the basic decision model; second, the specification of stake dependence; third, assumptions on the error term; and

[4]Advantages and potential drawbacks are discussed in Andersen et al. (2006a).

finally, a finite mixture regression approach which accounts for heterogeneity in behavior. At the end of this section we also briefly discuss some of the issues typically encountered when estimating finite mixture regression models.

2.1 The basic decision model

The basic model of decision under risk should be able to accommodate a wide range of different behaviors. Sign- and rank-dependent models, such as cumulative prospect theory (CPT), capture robust empirical phenomena such as nonlinear probability weighting and loss aversion (Starmer 2000). Therefore, the flexible approach of CPT, lends itself to describing risk taking behavior. According to CPT, an individual values any binary lottery $\mathcal{G}_g = (x_{1g}, p_g; x_{2g}),\ g \in \{1, \ldots, G\}$, where $|x_{1g}| > |x_{2g}|$, by

$$v\left(\mathcal{G}_g\right) = v\left(x_{1g}\right) w\left(p_g\right) + v\left(x_{2g}\right)\left(1 - w\left(p_g\right)\right). \tag{1}$$

The function $v(x)$ describes how monetary outcomes x are valued, whereas the function $w(p)$ assigns a subjective weight to every outcome probability p. The lottery's certainty equivalent $\hat{ce}_g$ can then be written as

$$\hat{ce}_g = v^{-1}\left[v\left(x_{1g}\right) w\left(p_g\right) + v\left(x_{2g}\right)\left(1 - w\left(p_g\right)\right)\right]. \tag{2}$$

In order to make CPT operational we have to assume specific functional forms for the value function $v(x)$ and the probability weighting function $w(p)$. A natural candidate for $v(x)$ is a sign-dependent power function

$$v(x) = \begin{cases} x^{\alpha} & \text{if } x \geq 0 \\ -(-x)^{\beta} & \text{otherwise,} \end{cases} \tag{3}$$

which can be conveniently interpreted and which has also turned out to be the best compromise between parsimony and goodness of fit in the context of prospect theory (Stott 2006). For this specification of the value function, a separate parameter of loss aversion, i.e. capturing that "losses loom larger than corresponding gains", is not identifiable in our data.[5] As Koebberling and Wakker (2005) point out, loss aversion should be interpreted as the difference between risk aversion with respect to mixed gambles, encompassing both gains and losses, and nonmixed gambles, confined to single-domain outcomes. Our lottery design comprises nonmixed gambles only and, therefore, the concept of loss aversion in this interpretation cannot be applied to our analysis.

A variety of functions for modeling probability weights $w(p)$ have been proposed in the literature (Quiggin 1982; Tversky and Kahneman 1992; Prelec

[5]To see this, assume $v(x) = -\lambda(-x)^{\beta}$ for losses. For nonmixed loss lotteries the parameter of loss aversion λ cancels out in the definition of the certainty equivalent ce: $-\lambda(-ce)^{\beta} = -\lambda(-x_1)^{\beta} w(p) - \lambda(-x_2)^{\beta}(1 - w(p))$ holds for any arbitrary value of λ.

1998). We use the linear-in-log-odds function[6] introduced by Goldstein and Einhorn (1987) and applied by Lattimore et al. (1992):

$$w(p) = \frac{\delta p^{\gamma}}{\delta p^{\gamma} + (1-p)^{\gamma}}, \ \delta \geq 0, \ \gamma \geq 0. \tag{4}$$

We favor this specification because it has proven to account well for individual heterogeneity (Wu et al. 2004)[7] and its parameters have an intuitively appealing interpretation: The parameter γ largely governs the slope of the curve, whereas the parameter δ largely governs its elevation. The smaller the value of γ, the more strongly S-shaped is the curve. Thus, this parameter can be interpreted as a measure of departure from rationality if linear weighting is considered to be the standard for rationality. The parameter δ can be viewed as indicator of optimism: The larger the value of δ, the more elevated is the curve, ceteris paribus, i.e. the higher is the weight put on any probability. Linear weighting is characterized by $\gamma = \delta = 1$. In a sign-dependent model, the parameters may take on different values for gains and for losses, yielding a total of six behavioral parameters to be estimated.

2.2 Stake dependence

In order to address our focal question of stake-size effects, we introduce a dummy variable HIGH into the basic decision model, such that HIGH = 1 if the lottery under consideration contains high-stake payoffs amounting to 65 CHN or more, and HIGH = 0 otherwise. Each one of the behavioral model parameters $\omega \in \{\alpha, \beta, \gamma', \delta'\}$, with γ' and δ' comprising the domain-specific parameters for the slope and the elevation of the probability weighting functions, is assumed to depend linearly on HIGH in the following fashion:

$$\omega = \omega_0 + \omega_{\text{HIGH}} \times \text{HIGH}, \tag{5}$$

with ω_0 representing the respective low-stake parameter. This step adds another six behavioral parameters to the set of model parameters.

If relative risk aversion indeed changes with stake size, at least one of the coefficients of the high-stake dummy HIGH should turn out to be significantly different from zero. If the estimates of α_{HIGH} or β_{HIGH} were significant, the present model would be mis-specified, as the power function, used for estimation, cannot account for changing relative risk aversion. In this case, an alternative specification of the value function that can accommodate changing relative risk aversion would be called for. In particular, if the valuation of

[6] A necessary and sufficient preference condition for this specification is presented in Gonzalez and Wu (1999).

[7] Moreover, the function generally fits equally well as the two-parameter function developed by Prelec (1998).

monetary outcomes is the driving force behind the observed change in risk tolerance over gains, α_{HIGH} should be negative, material in size and statistically significant.

2.3 Error specification

In the course of the experiment, risk taking behavior of individual $i \in \{1, \ldots, N\}$ was measured by her certainty equivalents ce_{ig} for a set of different lotteries $\mathcal{G}$. Since the behavioral model explains *deterministic* choice, an individual's actual certainty equivalents ce_{ig} are bound to deviate from the predicted certainty equivalents $\hat{ce}_g$ by an error ϵ_{ig}, i.e. $ce_{ig} = \hat{ce}_g + \epsilon_{ig}$. There may be different sources of error, such as carelessness, hurry or inattentiveness, resulting in accidentally wrong answers (Hey and Orme 1994). The Central Limit Theorem supports our assumption that errors are normally distributed and simply add white noise.

Furthermore, we allow for three different sources of heteroskedasticity in the error variance. First, for each lottery subjects have to consider 20 certain outcomes, which are equally spaced throughout the lottery's outcome range $|x_{1g} - x_{2g}|$. Since the observed certainty equivalent ce_{ig} is calculated as the arithmetic mean of the smallest certain amount preferred to the lottery and the subsequent certain amount, the error is proportional to the outcome range, which has to be taken account of by the estimation procedure.

Second, since our approach models a representative agent's behavior, an individual's choices will most likely depart from the average prediction. As subjects may be heterogeneous with respect to their previous knowledge, their ability of finding the correct certainty equivalent as well as their attention span, we expect the error variance to differ by individual. Rather than imposing additional assumptions on the error distribution[8] we estimate the standard deviations of the individual errors σ_{ig} directly by

$$\sigma_{ig} = \xi_i \left| x_{1g} - x_{2g} \right|, \tag{6}$$

where ξ_i is an individual-specific parameter. As expected, $\xi_i = \xi$ is rejected by a likelihood ratio test with a p-value close to zero, favoring specification of individual errors.

Third, lotteries in the gain domain may be evaluated differently from the ones in the loss domain.[9] Therefore, we additionally allow for domain-specific variance in the error term. Our assumption is justified by a likelihood ratio test

[8] As the graphs in Appendix B show, the distributions of the estimated ξ_i clearly depart from normality.

[9] As all the lotteries under consideration here are single-domain ones with a maximum of two non-zero outcomes only, there seems to be no need for modeling any lottery-specific errors, such as heteroskedasticity due to the position of the lottery in the sequence of choices, in addition to range dependence and domain dependence. The approach adopted in this paper is to characterize average behavior of large groups of people. Any sequence effects at the individual level disappear at the aggregate level if the order of the lotteries is individually randomized.

which rejects that ξ_i are the same for gains and losses. These considerations imply that we control for all three types of heteroskedasticity in the estimation procedure, which adds two more parameters per individual, i.e. 153 parameters for gains and 153 parameters for losses, to the econometric model.

2.4 Accounting for heterogeneity

A suitable estimation procedure, such as maximum likelihood, yields estimates for the average values of the behavioral parameters $\theta = (\alpha', \beta', \gamma', \delta')'$. If there is heterogeneity of a substantive kind, i.e. if there are several distinct data generating processes, a representative-agent model may be inferior to a model allowing for distinct types. For this reason, we estimate a finite mixture model which accounts for heterogeneity in a parsimonious way. The basic idea of the mixture model is assigning an individual's risk-taking choices to one of C different types of behavior, each characterized by a distinct vector of parameters $\theta_c = (\alpha_c', \beta_c', \gamma_c', \delta_c')'$, $c \in \{1, \ldots, C\}$. The estimation procedure yields estimates of the relative sizes of the different groups π_c, as well as the group-specific parameters θ_c of the underlying behavioral model.

In this paper we define groups across decision domains, i.e. each individual is classified on the basis of *all* of her choices over both gains and losses. In principle, one could analyze behavior over gains separately from behavior over losses. However, as we show in Appendix A, such an approach is inferior to an overall classification. Therefore, we define behavioral groups in a domain-independent way and estimate π_c jointly for gains and losses.[10]

Given our assumptions on the distribution of the error term, the density of type c for the i-th individual can be expressed as

$$f\big(ce_i, \mathcal{G}; \theta_c, \xi_i\big) = \prod_{g=1}^{G} \frac{1}{\sigma_{ig}} \phi \left(\frac{ce_{ig} - \hat{ce}_g\big(\mathcal{G}_g; \theta_c\big)}{\sigma_{ig}} \right), \tag{7}$$

where $\phi(\cdot)$ denotes the density of the standard normal distribution and ξ_i accounts for individual-specific heteroskedasticity. Since we do not know *a priori* which group a certain individual belongs to, the relative group sizes π_c are interpreted as probabilities of group membership. Therefore, each individual density of type c has to be weighted by its respective mixing proportion π_c, which is unknown and has to be estimated as well. Taking the sum over the weighted type-specific densities yields the individual's contribution to the model's likelihood function $L(\Psi; ce, \mathcal{G})$. The log likelihood of the finite mixture regression model is then given by

$$\ln L\,(\Psi; ce, \mathcal{G}) = \sum_{i=1}^{N} \ln \sum_{c=1}^{C} \pi_c \; f\big(ce_i, \mathcal{G}; \theta_c, \xi_i\big), \tag{8}$$

[10]The estimates of the model viewing types as domain specific are presented in Appendix A.

where the vector $\Psi = (\theta_1', \ldots, \theta_C', \pi_1, \ldots, \pi_{C-1}, \xi_1', \ldots, \xi_N')'$ summarizes the parameters to be estimated.

2.5 Estimation

In order to deal with the issues of non-linearity and multiple local maxima encountered when maximizing the likelihood of a finite mixture regression model (McLachlan and Peel 2000), the iterative Expectation Maximization (EM) algorithm is the method of choice here (Dempster et al. 1977). This algorithm also provides an additional feature: It calculates, by Bayesian updating in each iteration, an individual's posterior probability τ_{ic} of belonging to group c, given the fit of the model. These posterior probabilities τ_{ic} represent a particularly valuable result of the estimation procedure. Not only does the procedure endogenously assign each individual to a specific group, but it also supplies us with a method of judging classification quality. The τ_{ic} can be used to calculate a summary measure of ambiguity, such as the *Normalized Entropy Criterion NEC* (Celeux 1996), in order to gauge the extent of dubious assignments. If all the τ_{ic} of the final iteration are either close to zero or one, all the individuals are unambiguously assigned to one specific group and a low measure of entropy is observed.

Furthermore, entropy measures provide an additional criterion to discriminate between models with differing numbers of types. Since the finite mixture regression model is defined over a pre-specified number of groups, a criterion for assessing the correct number of groups is called for. Classical criteria, such as the Akaike Information Criterion AIC or the Bayesian Information Criterion BIC, trade off model parsimony against goodness of fit, but do not measure the ability of the mixture to provide well separated and nonoverlapping components, which, ultimately is the objective of estimating mixture models (Celeux 1996).

Various problems may be encountered when maximizing the likelihood function of a finite mixture regression model and, therefore, a customized estimation procedure has to be used, which can adequately deal with these problems. Details of the estimation procedure are discussed in Appendix C (see also Bruhin et al. 2007).

3 Results

In the following sections we investigate the stake-size sensitivity of observed risk taking behavior and present the estimates of the decision model assuming one homogeneous type of preferences. Furthermore, we show that substantive heterogeneity is present in our data and discuss the quality of the classification procedure as well as the number of distinct behavioral types identified in the data. Finally, we characterize these different types by their average behavioral parameters and discuss the effect of stake size on each group's behavior.

3.1 Aggregate behavior

Result 1 On average, observed behavior exhibits the fourfold pattern of risk attitudes, predicted by prospect theory, for both low-stake and high-stake outcomes. Stake-specific behavior is subject to a strong framing effect, however: When gains are at stake, relative risk aversion increases with stake size at almost all levels of probability. In the loss domain no such clear picture emerges.

Support In Fig. 2, observed risk taking behavior is summarized by the median relative risk premia $\mathrm{RRP} = (ev - ce)/|ev|$, where ev denotes the expected value of a lottery's payoff and ce stands for its certainty equivalent. $\mathrm{RRP} > 0$ indicates risk aversion, $\mathrm{RRP} < 0$ risk seeking, and $\mathrm{RRP} = 0$ risk neutrality. The light gray bars in Fig. 2 represent the observed median RRP for low-

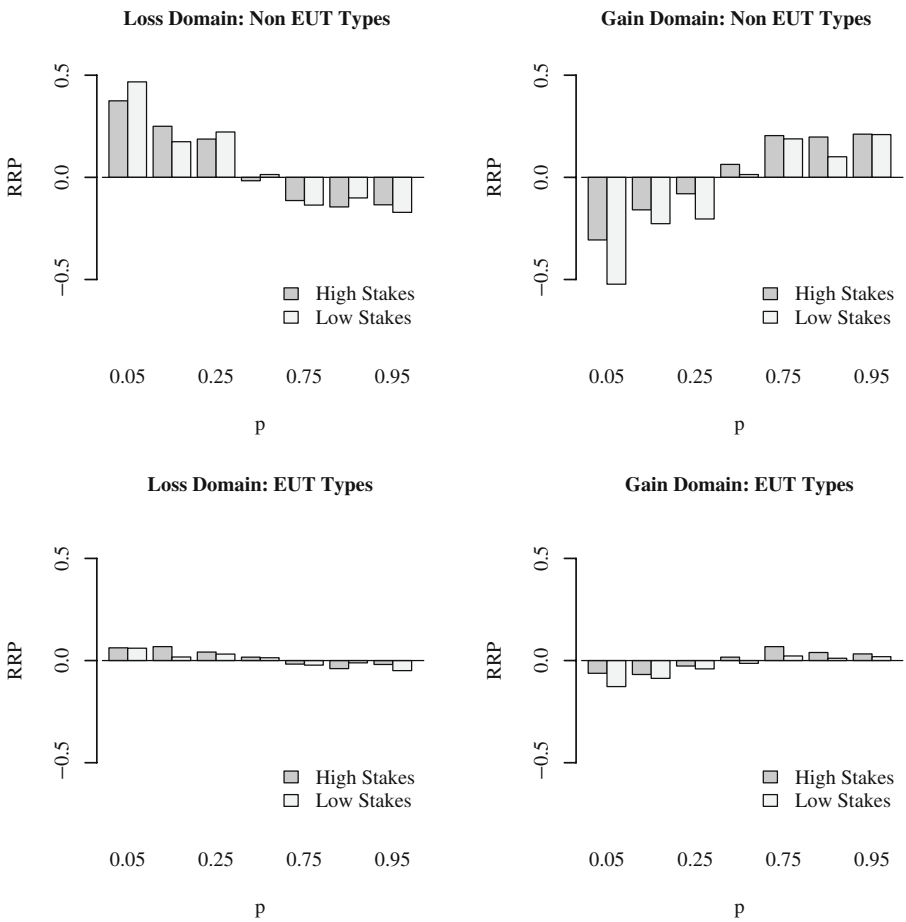


Fig. 2 Median relative risk premia by stake size

stake lotteries, the dark gray ones represent the respective high-stake median RRP. The median relative risk premia RRP, sorted by the probability p of the higher gain or loss, show a systematic relationship with p. For both low stakes and high stakes, subjects' choices display a fourfold pattern (Tversky and Kahneman 1992), i.e. they are risk averse for low-probability losses and high-probability gains, and they are risk seeking for low-probability gains and high-probability losses. Therefore, at first glance, average behavior is adequately described by a model such as CPT.

What the bar plots also reveal is that median relative risk premia differ substantially by stake level: When subjects' preferences exhibit increasing relative risk aversion, we should observe different low-stake and high-stake RRP, namely, high-stake choices should be relatively less risk tolerant than low-stake choices. Inspection of Fig. 2 confirms that, in the gain domain, median high-stake choices are considerably less risk seeking for small probabilities and somewhat more risk averse for large probabilities than their median low-stake counterparts. For losses, the evidence is not so clear-cut, however. At some levels of probability, low-stake median RRP display relatively higher risk aversion than high-stake RRP, and at some other levels the reverse is true.

In order to judge whether the distributions of the stake-dependent RRP are significantly different from each other, we performed a series of Wilcoxon signed-rank tests for each level of probability, which yield the following results at the conventional level of significance of 5%: With the exception of the probability of 95%, all the low-stake RRP over gains are significantly smaller than the high-stake ones. We therefore conclude that there is a significant stake effect in the data on choices over gains: On average, people are relatively more risk averse for high gains than for low gains.

In the loss domain, no consistent picture emerges: Low-stake RRP are significantly smaller at three levels of probability ($p \in \{0.10, 0.75, 0.95\}$), significantly larger at one level ($p = 0.05$), and insignificantly different at the remaining three levels of probability ($p \in \{0.25, 0.50, 0.90\}$). Therefore, we conclude that there is no obvious systematic relationship between stake-size effect and level of probability for loss lotteries.

Our data show behavior consistent with nonlinear probability weighting, but also a substantial framing effect. Relative risk aversion increases with stake size, albeit only for gains. When subjects evaluate the same lotteries framed as losses rather than as gains, their relative risk aversion does not systematically increase. In fact, no coherent pattern of stake-dependent behavior under losses emerges. This sensitivity to framing, already visible at the level of observed behavior, clearly excludes expected utility theory from the list of eligible models for describing average risk taking behavior.

We now turn to one of our major concerns, namely, whether the change in relative risk aversion over gains can be attributed to a specific component of lottery evaluation.

Result 2 In the aggregate model, the estimated curvatures of the value functions do not significantly change with rising stakes.

Table 2 Classification of behavior: overall behavioral types

	Gains				Losses		
	Pooled (1)	EUT (2)	Non-EUT (3)		Pooled (4)	EUT (5)	Non-EUT (6)
π		0.266	0.734	π		0.266	0.734
		(0.026)	(0.026)			(0.026)	(0.026)
α_0	0.467	0.996	0.430	β_0	1.165	1.157	1.177
	(0.109)	(0.136)	(0.116)		(0.110)	(0.136)	(0.120)
α_{HIGH}	0.047	−0.080	0.066	β_{HIGH}	−0.038	−0.137	−0.106
	(0.158)	(0.165)	(0.167)		(0.162)	(0.178)	(0.178)
γ_0	0.316	0.863	0.225	γ_0	0.383	0.802	0.284
	(0.012)	(0.067)	(0.013)		(0.012)	(0.067)	(0.012)
γ_{HIGH}	0.056	0.026	0.058	γ_{HIGH}	0.045	0.027	0.046
	(0.012)	(0.023)	(0.012)		(0.012)	(0.024)	(0.012)
δ_0	1.304	0.952	1.265	δ_0	0.913	0.912	0.917
	(0.076)	(0.094)	(0.080)		(0.052)	(0.090)	(0.058)
δ_{HIGH}	−0.324	−0.040	−0.344	δ_{HIGH}	0.070	0.106	0.099
	(0.095)	(0.090)	(0.098)		(0.077)	(0.091)	(0.087)
ln L	31,536	32,580					
Parameters	318	331					
Observations	8,560	8,560					

Standard errors in parentheses are based on the bootstrap with 2,000 replications. Parameter vectors include estimates of $\hat{\xi}_i$ for domain- and individual-specific error parameters

Support Table 2 contains the parameter estimates for the decision model discussed in Section 2. For the time being, we focus on the average parameter estimates displayed in columns (1) and (4), labeled "Pooled". The curvature parameters of the value functions over low stakes are denoted by α_0 for gains and β_0 for losses.[11] α_{HIGH} and β_{HIGH} represent the corresponding estimated coefficients of the high-stake dummy HIGH, measuring the change in curvature brought about by increased stake levels. For both domains, the estimates for α_{HIGH} and β_{HIGH} are small in size, and the bootstrapped standard errors, reported in parentheses below the respective point estimates, indicate that the coefficients are not significantly different from zero. Furthermore, when a restricted model with stake-invariant curvature parameters is estimated, the likelihood ratio test of the restricted model against the unrestricted one renders a p-value of 0.911. This test result implies that the hypothesis of equal curvatures over both ranges of outcomes cannot be rejected.

If the valuation of monetary outcomes were the carrier of increasing relative risk aversion over gains, the estimates of α_{HIGH} would have to be negative, statistically significant and, presumably, quite sizable, since the specification of the value function as a power function can only accommodate constant relative

[11]The curvature for low-stake gains is estimated to be 0.467, in line with numerous previous findings (Stott 2006). The estimate for β_0 amounts to 1.165, indicating slight concavity of the value function. However, its curvature is not statistically distinguishable for linearity. That the value function for losses is not convex, as predicted by prospect theory, is not an unusual finding (Abdellaoui 2000; Bruhin et al. 2007).

risk aversion. As the estimation results show, however, this is not the case. Therefore, we conclude that changing attitudes towards monetary outcomes are not responsible for the observed increase in relative risk aversion. This finding also holds for alternative specifications of the value function that are sufficiently flexible to capture changing relative risk aversion, such as the expo-power function introduced by Saha (1993).

As the curvature of the value function is robust to stake size, the observed increase in relative risk aversion over gains has to be driven by the other component of lottery evaluation, probability weighting, as the next result confirms.

Result 3 At the aggregate, high-gain probability weights deviate significantly from low-gain probability weights. No substantial difference in stake-dependent probability weights results for losses.

Support We first discuss our findings for the gain domain. A first indication of the stake sensitivity of probability weights for gains can already be found in the bar plots in Fig. 2. The *differences* in the observed stake-dependent RRP decrease markedly with increasing probability level, suggesting a substantial interaction effect.[12]

Inspection of column (1) of Table 2 confirms that the estimated change in the elevation of the curve, measured by δ_{HIGH}, is significantly negative and substantial in size, implying a major decrease in elevation from 1.304 to 0.913. Moreover, the change in the slope of the probability weighting function γ_{HIGH} is significantly positive (0.056), implying a slightly less strongly S-shaped curve for high stakes. The impact of these parameter changes on the shape of the probability weighting function can be examined in Fig. 3. The top panel of the figure shows, for each decision domain, the estimated probability weighting curves for low stakes, defined by $\text{HIGH} = 0$, plotted against the high-stake curves, defined by $\text{HIGH} = 1$. Evidently, the high-gain function is less elevated and slightly less strongly curved than the low-gain function, indicating a substantial decrease in probability weights.

However, significant changes in single parameter estimates do not tell the whole story. Since the probability weights are a nonlinear combination of two parameters, inference needs to be based on γ and δ jointly. Therefore, the percentile bootstrap method (using 2,000 replications) was employed to

[12]The study on risky gains by Kachelmeier and Shehata (1992), conducted in Beijing as well, finds that stake size interacts significantly with probability level, which is in line with our findings, but their data set is not sufficiently rich to draw any conclusions on relative contributions of outcome valuation and probability weighting. Furthermore, observed certainty equivalents in our data set show a clearly defined fourfold pattern of risk attitudes for both low stakes and high stakes, whereas Kachelmeier and Shehata find practically no risk aversion in choices over low stakes. The authors attribute this lack of risk aversion to the specifics of their elicitation procedure: Certainty equivalents were elicited as minimum selling prices, which seems to have induced some kind of loss aversion in subjects' responses.

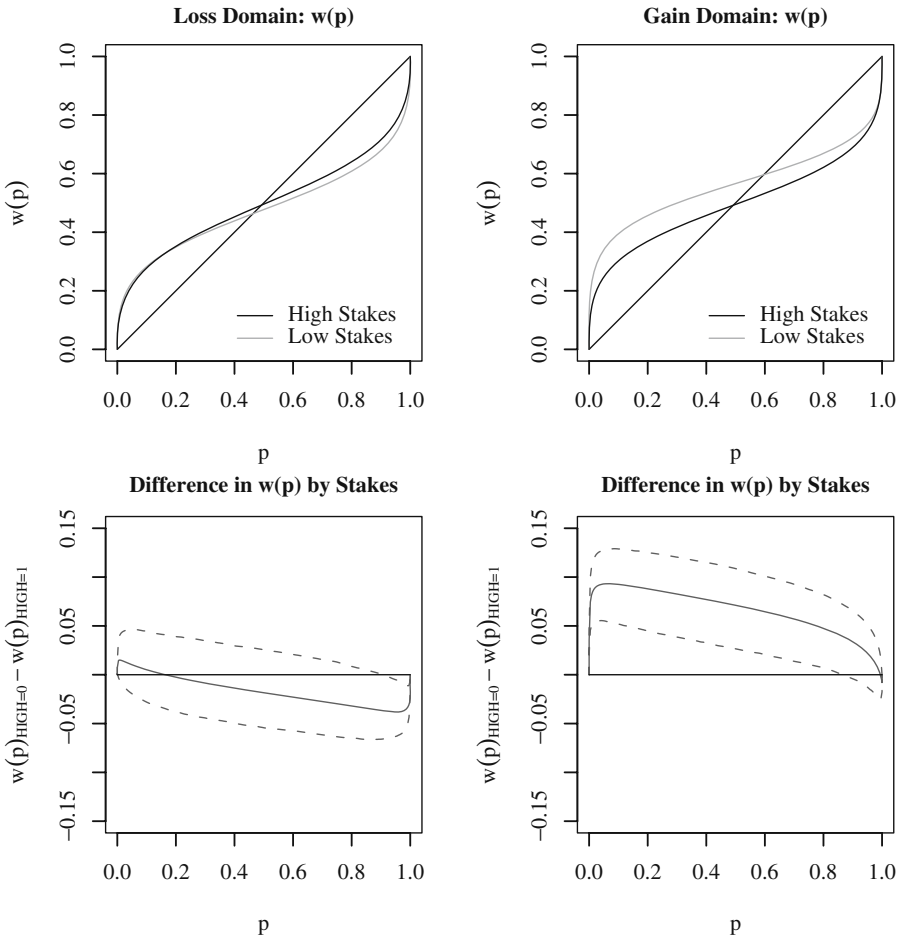


Fig. 3 Average probability weights by stake size. *Dashed lines*: 95%-confidence bands based on the percentile bootstrap method

construct the 95%-confidence bands for the difference in the low-stake and the high-stake probability weighting curves. In order to judge the overall effect of rising stakes on the shape of the probability weighting function, we inspect the bottom panel of Fig. 3, depicting the confidence bands for the stake-dependent differences in probability weights. Whenever a confidence band includes the zero line, the hypothesis of stake-invariant probability weights cannot be rejected. The graph on the left hand side for the gain domain, however, shows that the difference between low-gain probability weights and high-gain probability weights is indeed statistically significant over nearly the whole range of probabilities, mirroring our findings on the observed RRP. Therefore, we have conclusive evidence that the high-gain probability weighting curve departs significantly and substantially from the low-gain curve.

In the domain of losses, a totally different picture emerges. The top panel of Fig. 3 depicts practically overlapping low-loss and high-loss probability weighting curves. The high-loss curve is slightly less strongly S-shaped, which is also reflected in the significant parameter estimate for γ_{HIGH} in column (4) of Table 2, amounting to 0.045. However, this immaterial difference in the stake-dependent slope parameters does not imply a significant difference in the overall shape of the curves: The bottom panel of Fig. 3 shows that the 95%-confidence band for the difference in the stake-dependent probability weighting curves over losses includes the zero line practically for all levels of probability. This finding implies that, in choices framed as losses, stake effects are negligible, in line with the lack of any stake-dependent pattern diagnosed in the observed RRP.

Our findings demonstrate that probability weights are the carrier of changing risk tolerance observed in the domain of gains, and suggest that prospect theory, and for that matter many other decision theories which postulate stake-independent probability weighting, cannot adequately deal with risk taking choices involving major changes in stake levels.

3.2 Heterogeneous types of behavior

In the previous section we have only considered the evidence for the average decision maker. If there is heterogeneity in the population, in the sense that a single preference theory cannot adequately capture behavior, the parameter estimates of the pooled model may be misleading. For this reason, the analysis is extended to account for latent heterogeneity by estimating a finite mixture regression model.

So far we have not addressed the issue whether a finite mixture regression model is actually to be preferred over a single-component model in the first place, and what the number of groups C in the mixture model should be. In order to deal with these questions the researcher needs a criterion for assessing the correct number of mixture components. The literature on model selection in the context of mixture models is quite controversial, however, and there is no best solution.[13] For this reason, rather than relying on a single measure, we examine several criteria with differing characteristics to get a handle on the problem of model selection: the AIC, the BIC as well as the NEC. According to these criteria, the model which minimizes the respective criterion value should be preferred.

Result 4 There is substantive heterogeneity in individual risk preferences rendering the aggregate model inferior to a mixture of two distinct types of behavior.

[13] "The problem of identifying the number of classes is one of the issues in mixture modeling with the least satisfactory treatment." (Wedel 2002, p. 364).

Table 3 Model selection criteria

	AIC	BIC	NEC
$C = 1$	−62,436	−60,193	n.a.
$C = 2$	−64,498	−62,163	**0.0018**
$C = 3$	**−65,405**	**−62,978**	0.0035

Support We calculated AIC, BIC and NEC for three different model sizes, $C \in \{1, 2, 3\}$, presented in Table 3.

AIC and BIC are highest at $C = 1$, thus $C > 1$ is clearly favored over $C = 1$. As the NEC criterion is not defined for $C = 1$, Biernacki et al. (1999) argue in favor of a multi-component model if there is a $C > 1$ with $\mathrm{NEC}(C) \leq 1$, which clearly is the case here. Given the unanimous recommendation by all three criteria we conclude that a finite mixture model is superior to a representative-agent model.

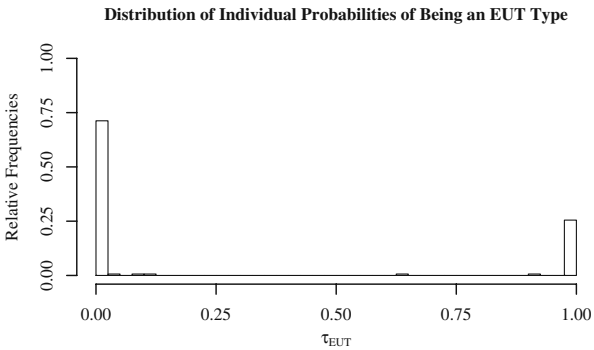
With regard to the choice between $C = 2$ and $C = 3$, the three-group classification seems to be favored by the classical criteria but not by NEC. From the point of view of entropy two groups are sufficient to characterize behavior. As entropy is generally extremely low for both the two-group and three-group classifications, both models seem quite sensible. However, it can be shown that for $C = 3$ the majority type gets divided into two different subtypes exhibiting qualitatively similar parameter estimates whereas the minority type, which is quite distinct, remains stable in size as well as group membership.[14] As the three-group classification does not offer additional material insights, we limit our discussion to $C = 2$.

The low value of NEC in our analysis indicates that nearly all the individuals can be unambiguously assigned to one of the two types. This clean segregation can also be inferred from the distributions of the posterior probabilities of group assignment in Fig. 4: τ_{EUT} denotes the posterior probability of belonging to the first group, which can be characterized, as we will demonstrate below, as expected utility maximizers ("EUT types"): The individuals' posterior probabilities of being an expected utility maximizer are either close to one or close to zero for practically all the individuals. The histogram also shows that the EUT group encompasses a minority of the decision makers, whereas the other group represents a majority of approximately 73% of the subjects.

The subsequent group of results addresses the focal questions: How can these two different types be characterized? And in which way do they react to rising stake levels?

[14]Results are available upon request.

Fig. 4 Clean segregation



3.2.1 Minority behavior

Result 5 The minority type, constituting about 27% of the subjects, can essentially be characterized as expected value maximizers over both low- and high-outcome ranges.

Support The estimates of the behavioral parameters for the minority type are displayed in columns (2) and (5) of Table 2. The relative group size of the minority type is estimated to be 0.266, matching the size of the corresponding bar in the histogram of Fig. 4.

In order to be able to characterize decisions as consistent with expected value maximization, both the value functions and the probability weighting functions are required to be linear. Turning to outcome valuation, we observe that α_0 and β_0 are not statistically distinguishable from one, as the standard errors reveal. Furthermore, the coefficients of the high-stake dummy are not significantly different from zero, indicating the robustness of the value function curvatures to increasing stake size. Therefore, we conclude that the value functions over both gains and losses are essentially linear and unresponsive to stake size.

Linearity of the second model component, probability weighting, holds if the parameter estimates for both γ and δ are equal to one. The low-stake parameter estimates for δ_0 in columns (2) and (5) are not distinguishable from one, but the respective ones for γ_0 are. However, inspection of the probability weighting curves in Fig. 5 confirms that departures from linear probability weighting are insubstantial. Furthermore, for both gains and losses, no stake-size effect is visible in slope or elevation of the probability weighting curves, as both γ_{HIGH} and δ_{HIGH} are insignificantly different from zero, and the 95%-confidence bands for the difference in the stake-dependent probability weighting curves include the zero line, as confirmed by the bottom panel of Fig. 5. These findings suggest that the minority type of decision makers behaves

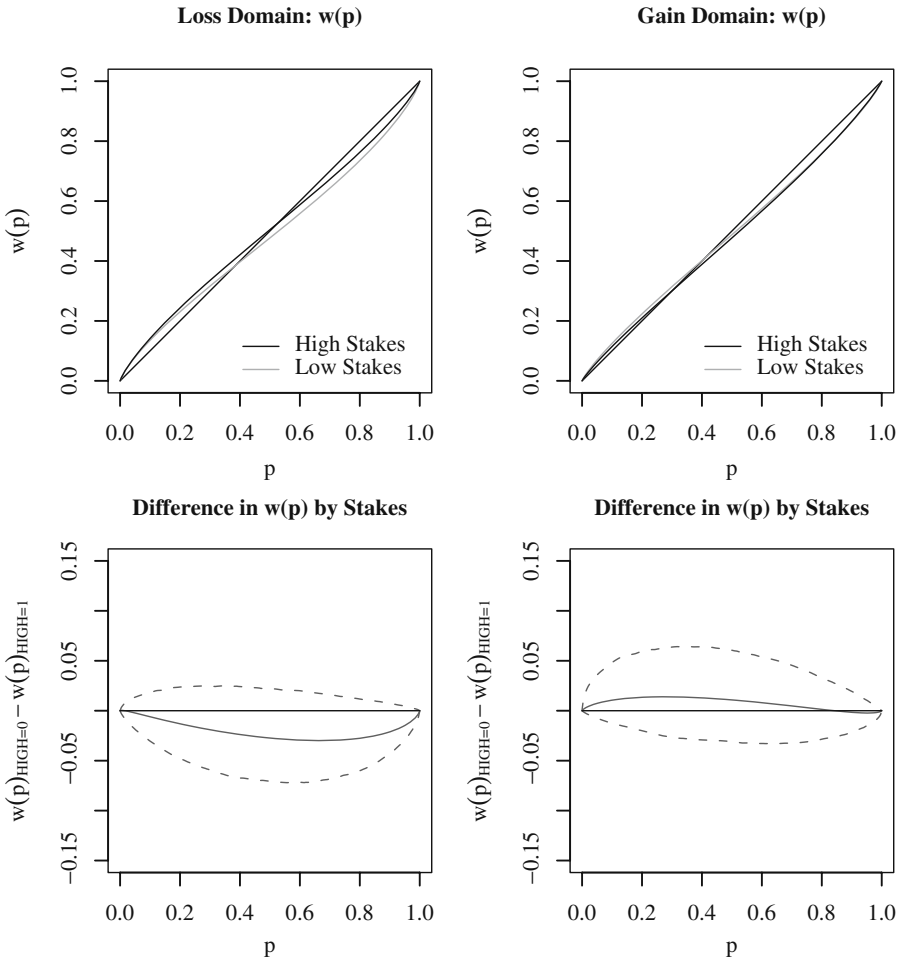


Fig. 5 Probability weights by stake size: EUT types. *Dashed lines*: 95%-confidence bands based on the percentile bootstrap method

essentially as expected value maximizers, and therefore consistently with EUT. These conclusions, based on the estimation results, also bear out at the level of observed behavior. The EUT types' median relative risk premia in the bottom panel of Fig. 6 are close to zero, indicating near risk neutrality for both low stakes and high stakes.

Obviously, the minority's behavior is robust over the whole outcome range and can, therefore, not account for increasing relative risk aversion observed in the aggregate data. As the next result shows, the second group of individuals, constituting approximately 73% of the subjects, exhibit a completely different set of behavioral parameter values.

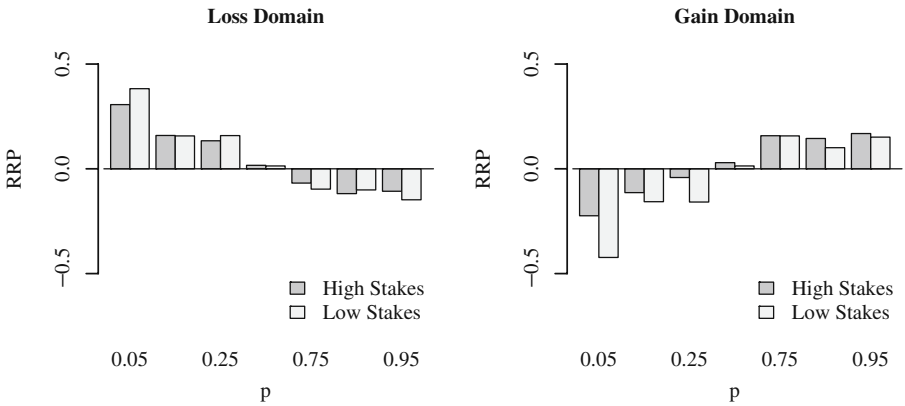


Fig. 6 Type-specific median relative risk premia by stake size

3.2.2 Majority behavior

Result 6 The majority group's behavior is characterized by nonlinear probability weighting. Whereas value function parameters remain stable over the whole outcome range in both decision domains, probability weights for gains do not. The low-gain probability weighting curve is characterized by significantly more optimistic weighting of probabilities than the high-gain curve. No such effect is present in the probability weighting curves for losses, however.

Support The majority group, labeled "Non-EUT", consists of about 73% of the individuals. As in the pooled model, value function parameter estimates of α_{HIGH} and β_{HIGH} are not significantly different from zero, as can be seen in columns (3) and (6) in Table 2. Again, the observed change in relative risk aversion over gains cannot be attributed to the valuation of monetary outcomes.

In order to examine the stake-specific probability weighting curves for the majority group we inspect the top panel of Fig. 7. Our findings on the majority group's curves reflect the same patterns of stake sensitivity as the pooled ones do: For both gains and losses the curves are inverted S-shaped, and there is a major domain-specific difference. In the loss domain the stake-specific curves practically coincide, and their difference is not statistically significant, as the left hand side of the bottom panel of Fig. 7 confirms. In the gain domain, however, we find the high-stake probability weighting curve to be substantially less elevated than the low-stake one. This change is brought about by the significant stake sensitivity of the elevation parameter over gains, reflected by the estimate for δ_{HIGH}, −0.344 (see column (3) in Table 2). The high-gain

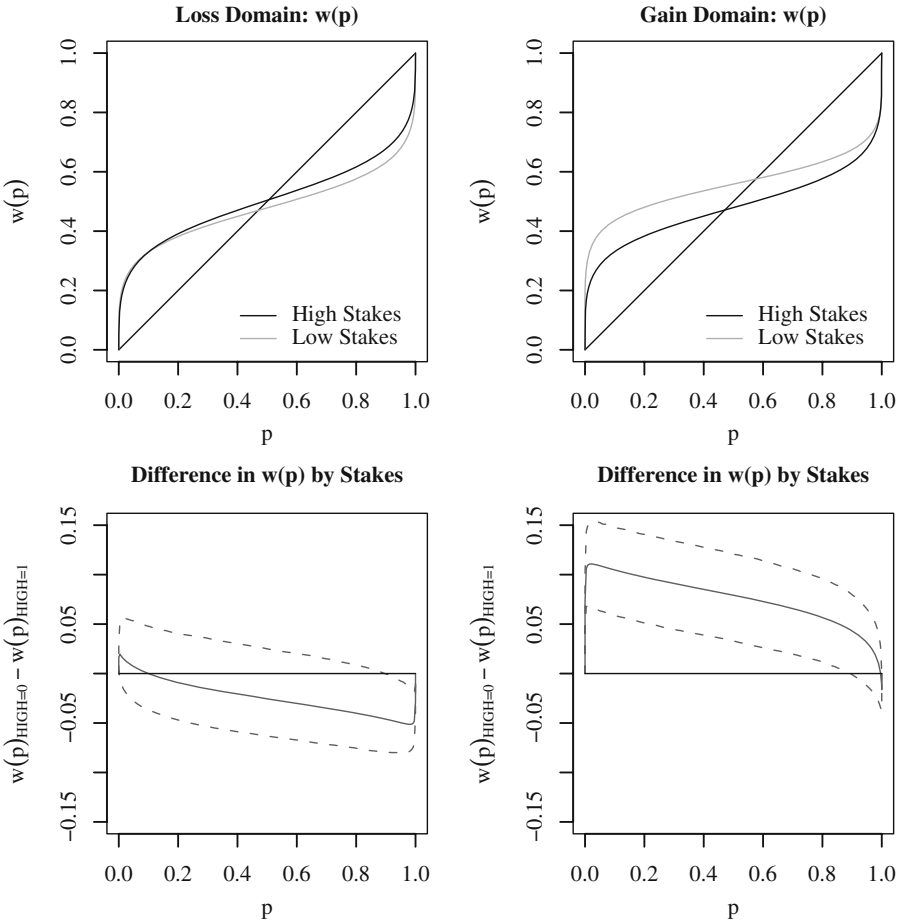


Fig. 7 Probability weights by stake size: non-EUT types. *Dashed lines*: 95%-confidence bands based on the percentile bootstrap method

probability weighting function is also slightly less curved than the low-gain one, as γ_{HIGH} is estimated to be 0.058.

The joint impact of these parameter changes is statistically significant, as the right hand side of the bottom panel of Fig. 7 shows. Therefore, we conclude that increasing relative risk aversion over gains is mainly attributable to the Non-EUT types' behavior who weight high-gain probabilities significantly and substantially less optimistically than low-gain ones. These effects can also be traced back in the pattern of observed choices: The top panel of Fig. 7 displays a substantial stake-dependent difference, particularly over smaller probabilities, in the Non-EUT types' median RRP, which is much

more pronounced than the respective difference in the pooled data, shown in Fig. 2.

The results of the finite mixture regression demonstrate that there is substantive heterogeneity in risk taking behavior, which may be glossed over when focusing on a single-preference model. Only one distinct group of individuals is prone to changing risk tolerance when stakes are increased. These Non-EUT types tend to evaluate low-gain prospects significantly more optimistically than high-gain prospects. Thus, prospect theory, even though designed to explain non-EUT behavior, cannot account for this change in relative risk aversion.

4 Discussion and conclusions

This paper pursues three goals. First, it studies the effect of substantial real payoffs, framed as gains and losses, on risk taking. Second, the paper analyzes the influence of rising stakes on the components of lottery evaluation, i.e. on the value and probability weighting functions. Third, it examines heterogeneity in risk taking behavior over varying stake levels. With regard to the first objective, we find a significant and sizable increase in relative risk aversion when gains are scaled up. In the domain of losses, however, no such clear effect is present in the data. We can only speculate on the potential reasons for this finding. One possible explanation is the use of different rules and heuristics when losses are at stake. This interpretation is supported by the empirical regularity that probability weights for gains differ systematically from probability weights for losses (Abdellaoui 2000; Bruhin et al. 2007). Another possibility is the hypothetical nature of the losses in our experiment. Here, losses are effectively gains and only appear as losses due to the framing of the decisions. Such an interpretation is consistent with the absence of a clear stake effect for purely hypothetical losses in Etchart-Vincent (2004). In any case, as subjects evaluate lotteries differently depending on the lotteries being framed as gains or as losses, expected utility theory is effectively ruled out as a valid description of behavior.

Concerning the second objective, we presented evidence that the increase in relative risk aversion over gains can be mainly attributed to a move of the average probability weighting function towards less optimistic weighting, whereas attitudes toward monetary payoffs remain essentially stable. That probability weights are the carrier of changing risk attitudes raises the question of the driving force behind this change. Unfortunately, little is known empirically about potential determinants of probability weights. However, evidence is accumulating that probability weights are systematically affected by specific characteristics of the decision situation. Not only do they vary with stake size but they are also sensitive to the delay of uncertainty resolution: Abdellaoui et al. (2009) find that departure of the probability weighting function from

linearity decreases with the delay of uncertainty resolution, resulting in an increase in risk tolerance over time (see also Noussair and Wu 2006).[15]

Recent generalizations of expected utility theory invoking emotions as rationale for probability weighting may offer a starting point for explaining the malleability of probability weights (Bell 1985; Gul 1991; Loomes and Sugden 1986; Wu 1999). Walther (2003), for instance, shows that an inverse S-shaped probability weighting function emerges endogenously from utility maximization when a decision maker's utility depends not only on monetary payoffs but also on elation and disappointment anticipated to materialize at uncertainty resolution. In particular, over- and underweighting of probabilities is driven by the balance of elation proneness and disappointment aversion: The stronger is disappointment aversion relative to elation proneness, the broader will be the range of pessimistically weighted gain probabilities and vice versa.

Such an affect-based approach may provide a unifying explanation for the observed stake and delay dependence of probability weights *if relative disappointment aversion significantly interacts with the context of the decision situation*.[16] In the case of stake dependence, it seems plausible that, when stakes increase substantially, anticipated feelings of disappointment get stronger relative to the elation one expects to feel when an advantageous low-probability event materializes. This effect would shift the probability weighting curve downwards and induce relatively higher risk aversion. With regard to uncertainty resolution the vividness of anticipated emotions might generally be lower for delayed lotteries than for immediately resolved ones, which would result in the probability weighting curve getting less strongly S-shaped.[17] To our knowledge, the theory of affective utility has not been directly tested empirically.[18] However, the study by Rottenstreich and Hsee (2001) may be interpreted as preliminary evidence for affect sensitivity of probability weights: The authors report that people tend to be less responsive to probabilities

[15]Similarly to our findings, the delay dependence of risk tolerance is solely due to a change in probability weights whereas delay has no discernible effect on the valuation of monetary outcomes.

[16]Such an interaction effect is not modeled by affective utility theories, as intensity of anticipated emotions is assumed to be independent of lottery characteristics.

[17]In principle, the framework of affective utility could also account for the absence of a stake effect for losses as well. Our Chinese subjects exhibit comparatively strong elation proneness over low-stake gains which diminishes considerably for high-stake gains. In the domain of low-stake losses elation seeking and disappointment aversion are more equally balanced in the first place and, therefore, there may be less leeway for decreasing elation proneness further. Note that the probability weighting curves for losses closely resemble the curve for high-stake gains. Of course, this conjecture is purely speculative.

[18]Abdellaoui and Bleichrodt (2007) tested Gul's theory of disappointment aversion and found it to be too parsimonious to explain their data.

when they react to emotion-laden targets, such as a kiss by one's favorite movie star or an electric shock, than they do in the case of comparatively pallid monetary outcomes. Moreover, van Winden et al. (2008) have recently shown that the intensity of hope and worry indeed decline with the delay of uncertainty resolution, supporting our conjecture.

Our results pose a number of potential problems to both theoretical and applied economics. As most theories of decision under risk typically assume separability of probability weights and outcome valuation, decision models may misrepresent risk preferences considerably when probability weights interact with the size of payoffs or other lottery characteristics in a material way. Our findings suggest that such an interaction with stake size is significant and substantial, which renders rank-dependent models, such as prospect theory, questionable when risk preferences over a wide range of outcomes are concerned. Models of affective utility, on the other hand, seem to be a promising point of departure but need to be extended to account for interaction effects as well.

In pursuing our third major objective, we demonstrated that there is substantive heterogeneity which can be parsimoniously characterized by two distinct behavioral types who either weight probabilities near linearly or nonlinearly with only the latter type exhibiting increasing relative risk aversion. Two clearly segregated groups of comparable size and characteristics were also found in two independent Swiss data sets (Bruhin et al. 2007) and, for choices over gains only, in a British data set (Conte et al. 2007), which suggests that this mix of preference types seems to be quite robust for the class of decisions studied here.

The question of heterogeneity is an important one in many economic contexts. For example, it drives the division of labor in organizations, the development of human capital and it may create strong selection of participants into particular markets. Moreover, heterogeneity may make the marginal individual quite different from the average one, which might be problematic for representative-agent models used in macroeconomics and finance (Camerer 2006; Cohen and Einav 2007). In addition, heterogeneity drives the market interactions of rational and boundedly rational agents. Theoretical and experimental work on market aggregation has shown that the mix of rational and irrational actors may be crucial for the aggregate outcome (Russell and Thaler 1985; Haltiwanger and Waldman 1985, 1989; Fehr and Tyran 2005, 2008). This literature demonstrates that whether individual mistakes would be erased or magnified depends on the nature of strategic interdependence. When behaviors are strategic complements, even a minority of players may be pivotal for the market outcome. Substantive heterogeneity is also important for economic policy. Regulatory policy should be designed in such a way that it creates large benefits for those who make errors, while imposing little or no harm on those who are fully rational. In order to be able to gauge a program's total net benefits the regulator needs to know the mix of behavioral

types. Therefore, knowing the composition of the population is important for designing cost efficient programs and regulations.

It is an open question whether nonlinear probability weighting should be considered as irrational. In Walther's model of affective utility nonlinear probability weighting is a consequence of utility maximization and, thus, not irrational *per se*. However, there is evidence that *anticipated* emotions deviate significantly from *experienced* emotions, which drives a wedge between decision utility and experienced utility, the ultimate standard for welfare judgments (Kahneman et al. 1997; van Winden et al. 2008). If nonlinear probability weights are manifestations of systematic errors utility maximization is based on the wrong premises and, consequently, behavior is merely boundedly rational. Therefore, investigation into the mechanisms underlying the malleability of probability weights should be high on the list of priorities for future research. Moreover, many real world decisions involve substantial risky payoffs as well as delayed resolution of uncertainty which may have countervailing effects on risk tolerance. New carefully designed experiments are called for that vary stake size and delay in order to derive meaningful parameter estimates which are useful for applied economics. Finally, future research needs to examine the robustness of our results with respect to other characteristics of the decision situation, such as complexity. It may well be the case that parameter estimates as well as the mix of behavioral types will change when decision tasks become more complex.

Acknowledgements We are particularly thankful to Mei Wang, Xiaofei Xie, Marc Schuerer, Ernst Fehr, George Loewenstein, and Roberto Weber, to the participants of the ESA World Conference 2007, the EEA-ESEM 2008 and the Behavioral Decision Research Seminar at Carnegie Mellon University, as well as to an anonymous referee. The usual disclaimer applies. This research was supported by the Swiss National Science Foundation (Grant 100012-109907).

Appendix A: Domain-specific behavior

The main goal of estimating a finite mixture model is to classify each individual to one of C different behavioral types. Our econometric model, as specified in Eq. 8, requires that each individual belongs to only one behavioral type, even though the behavioral parameters of each type are domain specific and may be different for decisions framed as gains or as losses. This makes the model very flexible in classifying the individuals into types which may exhibit domain-specific behavioral patterns, but retains the individual as unit of classification.

However, one can also estimate the finite mixture model separately for each domain. In such a case, behavioral types are only characterized within their specific domain and, consequently, each individual will be classified into two

types, one for decisions framed as gains and the other for decisions framed as losses. The (overall) log likelihood of such a specification,

$$\ln \tilde{L}\left(\tilde{\Psi}; ce, \mathcal{G}\right) = \left[\sum_{i=1}^{N} \ln \sum_{c=1}^{C} \tilde{\pi}_c^{(w)} f\left(ce_i^{(w)}, \mathcal{G}^{(w)}; \tilde{\theta}_c^{(w)}, \tilde{\xi}_i^{(w)}\right)\right] + \left[\sum_{i=1}^{N} \ln \sum_{c=1}^{C} \tilde{\pi}_c^{(l)} f\left(ce_i^{(l)}, \mathcal{G}^{(l)}; \tilde{\theta}_c^{(l)}, \tilde{\xi}_i^{(l)}\right)\right], \tag{9}$$

is completely separable into decisions framed as gains, (w), and decisions framed as losses, (l). Note that since the above model (9) does not nest the original model (8) we cannot test these two specifications against each other. Nevertheless, as Table 4 shows, the separated model achieves a lower log likelihood than the original model, the estimates of which are displayed in Table 2, even though it has one more parameter. It also results in a higher value of the BIC (−62,058 vs. −62,163). This shows that the above specification is clearly inferior to domain-independent typing.

At the level of individual type assignment, the main difference between the domain-specific model and the domain-independent one is that 13 people, identified as overall EUT type, move to the Non-EUT group for gains, which explains the lower percentage of EUT types in the domain-specific model (19.5% vs. 26.6%).

Table 4 Classification of behavior: domain-dependent types

	Gains				Losses		
	Pooled (1)	EUT (2)	Non-EUT (3)		Pooled (4)	EUT (5)	Non-EUT (6)
π		0.195	0.805	π		0.275	0.725
		(0.067)	(0.067)			(0.032)	(0.032)
α_0	0.467	1.028	0.428	β_0	1.165	1.165	1.174
	(0.109)	(0.178)	(0.115)		(0.110)	(0.134)	(0.121)
α_{HIGH}	0.047	−0.080	0.078	β_{HIGH}	−0.038	−0.124	−0.112
	(0.158)	(0.155)	(0.168)		(0.162)	(0.166)	(0.175)
γ_0	0.316	0.913	0.241	γ_0	0.383	0.782	0.279
	(0.012)	(0.120)	(0.027)		(0.012)	(0.076)	(0.015)
γ_{HIGH}	0.056	0.019	0.058	γ_{HIGH}	0.045	0.032	0.046
	(0.012)	(0.023)	(0.013)		(0.012)	(0.025)	(0.012)
δ_0	1.304	0.914	1.287	δ_0	0.913	0.906	0.915
	(0.076)	(0.147)	(0.087)		(0.052)	(0.094)	(0.059)
δ_{HIGH}	−0.324	−0.018	−0.357	δ_{HIGH}	0.070	0.103	0.101
	(0.095)	(0.106)	(0.097)		(0.077)	(0.087)	(0.085)
$\ln L$	31,536	16,132			31,536	16,401	
Parameters	318	166			318	166	
Observations	8,560	4,278			8,560	4,282	

Standard errors in parentheses are based on the bootstrap with 2,000 replications. Parameter vectors include estimates of $\hat{\xi}_i$ for domain- and individual-specific error parameters

Appendix B: Error parameter distribution

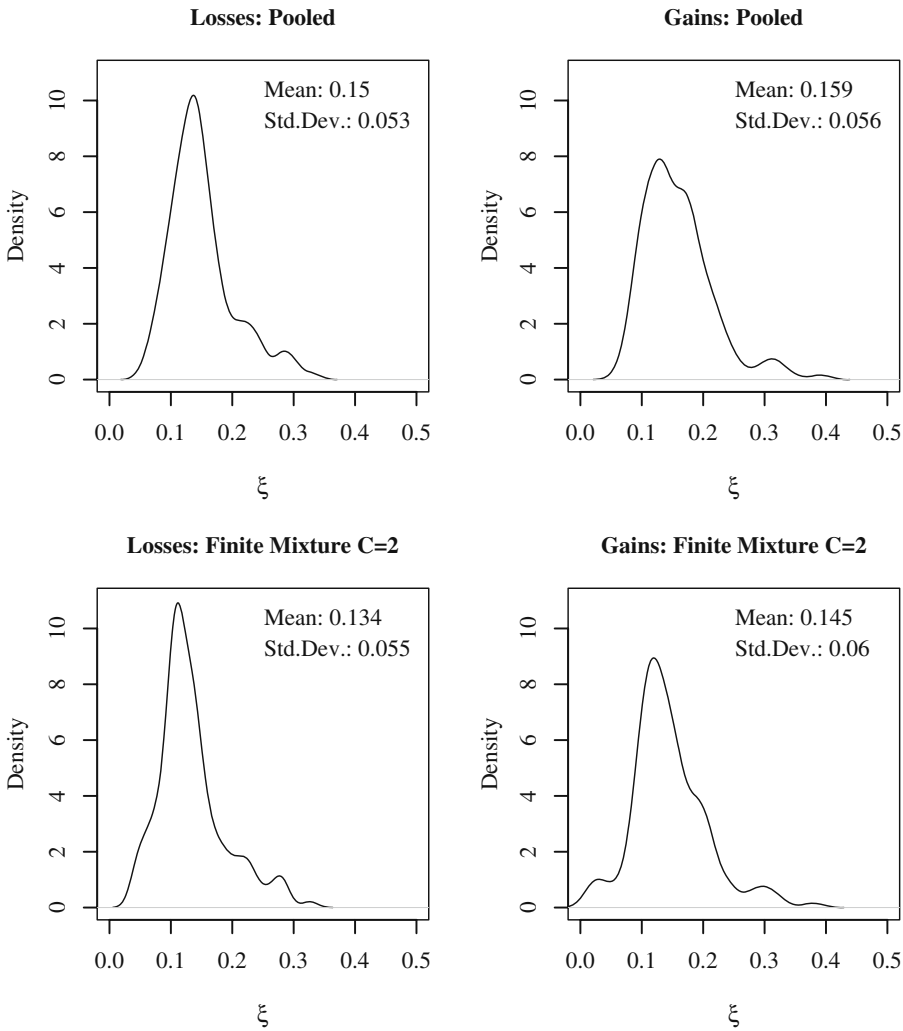


Fig. 8 Distributions of estimated $\hat{\xi}_i$ by domain

Appendix C: Estimation of the finite mixture regression model

As it is generally the case in finite mixture models, direct maximization of the log likelihood function

$$\ln L\big(\Psi; ce, \mathcal{G}\big) = \sum_{i=1}^{N} \ln \sum_{c=1}^{C} \pi_c \, f\big(ce_i, \mathcal{G}; \theta_c, \xi_i\big)$$

may encounter several problems, even if it is in principle feasible (for a general treatise see for example McLachlan and Peel (2000)). First, the highly non-linear form of the log likelihood causes the optimization algorithm to be rather slow or even incapable of finding the maximum. Second, the likelihood of a finite mixture model is often multimodal and, therefore, we have no guaranty that a standard optimization routine will converge towards the global maximum rather than to one of the local maxima.

However, if individual group-membership were observable and indicated by $t_{ic} \in \{0, 1\}$ the individual contribution to the likelihood function would be given by

$$\tilde{\ell}(\Psi_i; ce_i, \mathcal{G}, t_i) = \prod_{c=1}^{C} \left[\pi_c \, f(ce_i, \mathcal{G}; \theta_c, \xi_i)\right]^{t_{ic}}$$

By using the above formulation and taking logarithms, the complete-data log likelihood function

$$\ln \tilde{L}(\Psi; ce, \mathcal{G}, t) = \sum_{i=1}^{N} \sum_{c=1}^{C} t_{ic} \left[\ln \pi_c + \ln f(ce_i, \mathcal{G}; \theta_c, \xi_i)\right]$$

would follow directly. As relative group sizes sum up to one, their maximum likelihood estimates, $\hat{\pi}_c = 1/N \sum_{i=1}^{N} t_{ic}$, would be given analytically by the relative number of individuals in the respective group. Furthermore, the maximum likelihood estimates of the group-specific parameters could be obtained separately in each group by numerically maximizing the corresponding joint density function which would simplify the optimization problem considerably.

The EM algorithm proceeds iteratively in two steps, E and M, while it treats the unobservable t_{ic} as missing data. In the E-step of the $(k+1)$-th iteration the expectation of the complete-data log likelihood $\tilde{L}$, given the actual fit of the data $\Psi^{(k)}$, is computed. This yields, according to Bayes' law, the posterior probabilities of individual group-membership

$$\tau_{ic}\left(ce_i, \mathcal{G}; \Psi_i^{(k)}\right) = \frac{\pi_c^{(k)} \, f\left(ce_i, \mathcal{G}; \theta_c^{(k)}, \xi_i^{(k)}\right)}{\sum_{m=1}^{C} \pi_m^{(k)} \, f\left(ce_i, \mathcal{G}; \theta_m^{(k)}, \xi_i^{(k)}\right)}$$

which replace the unknown indicators of individual group-membership, t_{ic}.

Given $\tau_{ic}\left(ce_i, \mathcal{G}; \Psi_i^{(k)}\right)$, the complete-data log likelihood, $\tilde{L}$, is maximized in the following M-step which yields the updates of the model parameters,

$$\pi_c^{(k+1)} = \frac{1}{N} \sum_{i=1}^{N} \tau_{ic}\left(ce_i, \mathcal{G}; \Psi_i^{(k)}\right),$$

and

$$\left(\theta_1^{(k+1)}, \ldots, \theta_C^{(k+1)}, \xi_1^{(k+1)}, \ldots, \xi_N^{(k+1)}\right) = \underset{\theta_1,\ldots,\theta_C,\xi_1,\ldots,\xi_N}{\arg\max} \sum_{i=1}^{N} \sum_{m=1}^{C} \tau_{im}\left(ce_i, \mathcal{G}; \Psi_i^{(k)}\right) \times \ln f\left(ce_i, \mathcal{G}; \theta_m^{(k)}, \xi_i^{(k)}\right).$$

As Dempster et al. (1977) show, the likelihood never decreases from one iteration to the next, i.e. $L\left(\Psi^{(k+1)}; ce, \mathcal{G}\right) \geq L\left(\Psi^{(k)}; ce, \mathcal{G}\right)$, which makes the EM algorithm converge monotonically towards the nearest maximum of the likelihood function regardless whether this maximum is global or just local. Therefore, one may apply a stochastic extension, the Simulated Annealing Expectation Maximization (SAEM) algorithm proposed by Celeux et al. (1996), in order to overcome the EM algorithm's tendency to converge towards local maxima. In each iteration, there is a non-zero probability that the SAEM algorithm leaves the current optimization path and starts over in a different region of the likelihood function, which results in much higher chances of finding the global maximum. But this robustness against multimodality of the objective function comes at the cost of much higher computational demands.

As the EM algorithm is computationally highly demanding, even in its basic form, and tends to become tediously slow when close to convergence, our estimation routine relies on a hybrid estimation algorithm (Render and Walker 1984): It first uses either the EM or the SAEM algorithm and takes advantage of their robustness before it switches to the direct maximization of the log likelihood by the much faster BFGS algorithm.[19] The estimation routine in this form turned out to be efficient and robust as it reliably converged towards the same maximum likelihood estimates regardless of the randomly chosen start values.

References

Abdellaoui, M. (2000). Parameter-free elicitation of utilities and probability weighting functions. *Management Science, 46*, 1497–1512.

Abdellaoui, M., & Bleichrodt, H. (2007). Eliciting Gul's theory of disappointment aversion by the tradeoff method. *Journal of Economic Psychology, 28*, 631–645.

Abdellaoui, M., Diecidue, E., & Onculer, A. (2009). *Eliciting the impact of delayed resolution of uncertainty on attitude toward risk*. Unpublished manuscript.

Andersen, S., Harrison, G., Lau, M., & Rutstroem, E. (2006a). Elicitation using multiple price list formats. *Experimental Economics, 9*, 383–405.

Andersen, S., Harrison, G., Lau, M., & Rutstroem, E. (2006b). *Risk aversion in game shows*. University of Durham, Working Paper in Economics and Finance, 06/07.

Astebro, T., Mata, J., & Santos-Pinto, L. (2009). *Preference for skew in lotteries: Laboratory evidence and applications*. Unpublished manuscript.

Baltussen, G., Post, T., & van den Assem, M. (2008). *Risky choice and the relative size of stakes*. Unpublished manuscript.

[19]The procedure is written in the *R* environment (R Development Core Team 2006).

Beattie, J., & Loomes, G. (1997). The impact of incentives upon risky choice experiments. *Journal of Risk and Uncertainty, 14*, 155–168.

Bell, D. (1985). Disappointment in decision making under uncertainty. *Operations Research, 33*, 1–27.

Biernacki, C., Celeux, G., & Govaert, G. (1999). An improvement of the NEC criterion for assessing the number of clusters in a mixture model. *Pattern Recognition Letters, 20*, 267–272.

Binswanger, H. P. (1981). Attitudes toward risk: Theoretical implications of an experiment in rural India. *Economic Journal, 91*, 867–890.

Bombardini, M., & Trebbi, F. (2005). *Risk aversion and expected utility theory: A field experiment with large and small stakes*. Unpublished manuscript.

Bosch-Domenech, A., & Silvestre, J. (1999). Does risk aversion or attraction depend on income? An experiment. *Economics Letters, 65*, 265–273.

Bruhin, A., Fehr-Duda, H., & Epper, T. F. (2007). *Risk and rationality: Uncovering heterogeneity in probability distortion*. Socioeconomic Institute, University of Zurich, Working Paper, 0705.

Camerer, C. (1991). Recent tests of generalizations of expected utility theory. In W. Edwards (Ed.), *Utility: Measurement, theory and applications*. Amsterdam: Kluwer.

Camerer, C. (2006). *Behavioral economics*. Unpublished manuscript.

Camerer, C., & Hogarth, R. (1999). The effects of financial incentives in experiments: A review and capital–labor–production framework. *Journal of Risk and Uncertainty, 19*(1–3), 7–42.

Camerer, C., Issacharoff, S., Loewenstein, G., O'Donogue, T., & Rabin, M. (2003). Regulation for conservatives: Behavioral economics and the case for "asymmetric paternalism". *University of Pennsylvania Law Review, 151*, 1211–1254.

Camerer, C. F., & Fehr, E. (2006). When does economic man dominate social behavior? *Science, 311*.

Celeux, G. (1996). An entropy criterion for assessing the number of clusters in a mixture model. *Journal of Classification, 13*, 195–212.

Celeux, G., Chauveau, D., & Diebolt, J. (1996). Stochastic versions of the EM algorithm: An experimental study in the mixture case. *Journal of Statistical Computation and Simulation, 55*(4), 287–314.

Cohen, A., & Einav, L. (2007). Estimating risk preferences from deductible choice. *American Economic Review, 97*(3), 745–788.

Conte, A., Hey, J. D., & Moffat, P. G. (2007). *Mixture models of choice under risk*. Department of Economics and Related Studies, University of York, Discussion Paper, 2007/06.

Dempster, A., Laird, N., & Rubin, D. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B, 39*(1), 1–38.

El-Gamal, M. A., & Grether, D. M. (1995). Are people Bayesian? Uncovering behavioral strategies. *Journal of the American Statistical Association, 90*(432), 1137–1145.

Etchart-Vincent, N. (2004). Is probability weighting sensitive to the magnitude of consequences? An experimental investigation on losses. *Journal of Risk and Uncertainty, 28*(3), 217–235.

Fehr, E., & Tyran, J.-R. (2005). Individual irrationality and aggregate outcomes. *Journal of Economic Perspectives, 19*(4), 43–66.

Fehr, E., & Tyran, J.-R. (2008). Limited rationality and strategic interaction: The impact of the strategic environment on nominal inertia. *Econometrica, 76*(2), 353–394.

Goldstein, W., & Einhorn, H. (1987). Expression theory and the preference reversal phenomena. *Psychological Review, 94*, 236–254.

Gonzalez, R., & Wu, G. (1999). On the shape of the probability weighting function. *Cognitive Psychology, 38*, 129–166.

Grisley, W., & Kellog, E. (1987). Risk-taking preferences of farmers in northern Thailand: Measurements and implications. *Agricultural Economics, 1*, 127–142.

Gul, F. (1991). A theory of disappointment aversion. *Econometrica, 59*(3), 667–686.

Haltiwanger, J. C., & Waldman, M. (1985). Rational expectations and the limits of rationality: An analysis of heterogeneity. *American Economic Review, 75*(3), 326–340.

Haltiwanger, J. C., & Waldman, M. (1989). Limited rationality and strategic complements: The implications for macroeconomics. *Quarterly Journal of Economics, 104*(3), 463–483.

Hey, J. D., & Orme, C. (1994). Investigating generalizations of expected utility theory using experimental data. *Econometrica, 62*(6), 1291–1326.

Hogarth, R. M., & Einhorn, H. J. (1990). Venture theory: A model of decision weights. *Management Science, 36*(7), 780–803.
Holt, C. A., & Laury, S. K. (2002). Risk aversion and incentive effects. *American Economic Review, 92*(5), 1644–1655.
Holt, C. A., & Laury, S. K. (2005). Risk aversion and incentive effects: New data without order effects. *American Economic Review, 95*(3), 902–904.
Houser, D., Keane, M., & McCabe, K. (2004). Behavior in a dynamic decision problem: An analysis of experimental evidence using a Bayesian type classification algorithm. *Econometrica, 72*(3), 781–822.
Kachelmeier, S. J., & Shehata, M. (1992). Examining risk preferences under high monetary incentives: Experimental evidence from the People's Republic of China. *American Economic Review, 82*(5), 1120–1141.
Kahneman, D., Wakker, P., & Sarin, R. (1997). Back to Bentham? Explorations of experienced utility. *Quarterly Journal of Economics, 112*(2), 375–405.
Koebberling, V., & Wakker, P. (2005). An index of loss aversion. *Journal of Economic Theory, 122*, 119–131.
Kuehberger, A., Schulte-Mecklenbeck, M., & Perner, J. (1999). The effects of framing, reflection, probability, and payoff on risk preference in choice tasks. *Organizational Behavior and Human Decision Processes, 78*(3), 204–231.
Lattimore, P. K., Baker, J. R., & Witte, A. D. (1992). The influence of probability on risky choice. *Journal of Economic Behavior and Organization, 17*, 377–400.
Loomes, G., & Sugden, R. (1986). Disappointment and dynamic consistency in choice under uncertainty. *Review of Economic Studies, 53*(173), 271–282.
Markowitz, H. (1952). The utility of wealth. *The Journal of Political Economy, 60*(2), 151–158.
McLachlan, G., & Peel, D. (2000). *Finite mixture models. Wiley series in probabilities and statistics*. New York: Wiley-Interscience.
Noussair, C., & Wu, P. (2006). Risk tolerance in the present and the future: An experimental study. *Managerial and Decision Economics*, 27, 401–412.
Post, T., van den Assem, M., Baltussen, G., & Thaler, R. (2008). Deal or no deal? Decision making under risk in a large-payoff game show. *American Economic Review, 98*(1), 38–71.
Prelec, D. (1998). The probability weighting function. *Econometrica, 3*, 497–527.
Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior and Organization, 3*, 323–343.
R Development Core Team (2006). *R: A language and environment for statistical computing. R Foundation for statistical computing (Vienna, Austria)*. ISBN 3-900051-07-0. http://www.R-project.org.
Render, R. A., & Walker, X. (1984). Mixture densities, maximum likelihood and the EM algorithm. *SIAM Review, 26*(2), 195–239.
Rottenstreich, Y., & Hsee, C. K. (2001). Money, kisses, and electric shocks: On the affective psychology of risk. *Psychological Sciences, 12*(3), 185–190.
Russell, T., & Thaler, R. (1985). The relevance of quasi rationality in competitive markets. *American Economic Review, 75*(5), 1071–1082.
Saha, A. (1993). Expo-power utility: A 'flexible' form for absolute and relative risk aversion. *American Journal of Agricultural Economics, 75*, 905–913.
Smith, V., & Walker, J. (1993). Monetary rewards and decision cost in experimental economics. *Economic Inquiry, 31*, 245–261.
Stahl, D. O., & Wilson, P. W. (1995). On players' models of other players: Theory and experimental evidence. *Games and Economic Behavior, 10*, 218–254.
Starmer, C. (2000). Developments in non-expected utility theory: The hunt for a descriptive theory of choice under risk. *Journal of Economic Literature, 38*(2), 332–382.
Stott, H. P. (2006). Cumulative prospect theory's functional menagerie. *Journal of Risk and Uncertainty, 32*, 101–130.
Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty, 5*, 297–323.
van Winden, F., Krawczyk, M., & Hopfensitz, A. (2008). *Investment, resolution of risk, and the role of affect*. Tinbergen Institute Discussion Paper, TI 2008-047/1.

Walther, H. (2003). Normal-randomness expected utility, time preference and emotional distortions. *Journal of Economic Behavior and Organization, 52*(2), 253–266.

Weber, B. J., & Chapman, G. B. (2005). Playing for peanuts: Why is risk seeking more common for low-stakes gambles? *Organizational Behavior and Human Decision Processes, 97*, 31–46.

Wedel, M. (2002). Concomitant variables in finite mixture models. *Statistica Neerlandica, 56*(3), 362–375.

Wik, M., Kebede, T., Bergland, O., & Holden, S. (2004). On the measurement of risk aversion from experimental data. *Applied Economics, 36*, 2443–2451.

Wilcox, N. (1993). Lottery choice: Incentives, complexity and decision time. *The Economic Journal, 103*, 1397–1417.

Wu, G. (1999). Anxiety and decision making with delayed resolution of uncertainty. *Theory and Decision, 46*(2), 159–198.

Wu, G., Zhang, J., & Gonzalez, R. (2004). Decision under risk. In D. Koehler, & N. Harvey (Eds.), *The Blackwell handbook of judgment and decision making*. Oxford: Oxford University Press.