Rationality on the Rise: Why Relative Risk Aversion Increases with Stake Size

*

Helga Fehr-Duda   Adrian Bruhin   Thomas F. Epper

Renate Schubert

March 4, 2008

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 increase in relative risk aversion over gains cannot be captured by the curvature of the utility function. It is driven predominantly by a change in probability weighting of a majority group of individuals who exhibit more rational probability weighting at high stakes. 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.

*Authors' affiliation: Chair of Economics, Swiss Federal Institute of Technology Zurich, Switzerland. Corresponding Author: Helga Fehr-Duda, fehr@econ.gess.ethz.ch

1 Introduction

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 high stakes on risk tolerance has been debated since the early days of expected utility theory. In a seminal paper, Markowitz (1952) argued that risk preferences are likely to reverse from risk seeking over very small stakes to risk aversion over high stakes. While Markowitz did not test this conjecture experimentally, there is evidence by now that relative risk aversion is indeed greater when substantial amounts of money are at stake (Binswanger, 1981; Kachelmeier and Shehata, 1992; Holt and Laury, 2002), at least on average. 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 this change is actually a consequence of the way people value low versus high amounts of money, or whether some other component of lottery evaluation, such as probability weighting, is the driving force. Moreover, results based on aggregate data may gloss over potentially important differences in individual behavior.

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. The lotteries presented to the subjects were framed as gains and 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, to account for the existence of heterogeneous preference types, we used 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, Keane, and McCabe, 2004).

The following results emerge from our analysis. First, we find a strong and significant framing effect in 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 subject's behavior over identical lotteries framed as losses.

Second, 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 deviates less strongly from rational linear weighting than the respective function for low stakes. This change is particularly pronounced over the range of smaller probabilities, entailing less optimistic lottery evaluation at high stakes and, thus, increasing relative risk aversion. In the loss domain, however, no such change in probability weights can be inferred from the data.

Third, when allowing for heterogeneity of preference types, 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 at the aggregate level. Since it is the probability weights that are responsible for the change in relative risk aversion, using more flexible utility functions cannot adequately solve the problem of modelling increasing risk aversion.

While the observed framing effect lends support to sign-dependent decision models, such as prospect theory, stake dependence of probability weights, however, calls theories based on stake-invariant probability weighting 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), is substantive in the sense that a single preference model is unable to adequately describe behavior. This heterogeneity may render policy recommendations based on average parameter estimates inappropriate. Moreover, the mix of different types may be decisive for aggregate outcomes, as the literature on the role of bounded rationality under strategic complementarity and substitutability has shown (Haltiwanger and Waldman, 1985, 1989; Fehr and Tyran, 2005). The researcher will, therefore, need to deal with the potential stake sensitivity of the pivotal group.

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. Previous studies have focused on gains, on hypothetical payoffs or on quite limited payoff ranges, and have usually not addressed the issue of probability weighting (Hogarth and Einhorn, 1990; Bosch-Domenech and Silvestre, 1999; Kuehberger, Schulte-Mecklenbeck, and Perner, 1999; Weber and Chapman, 2005). Tversky and Kahneman (1992) suspected that probability weights might be sensitive to the level of outcomes, but they questioned whether increasing the complexity of decision theory was worthwhile the costs associated with such an endeavor. Some preliminary findings indeed suggest that probability weights may be stake dependent (Kachelmeier and Shehata, 1992; Camerer, 1991), but so far no conclusive evidence concerning the size and the importance of this effect has been presented. Moreover, findings based on hypothetical payoffs may be of limited relevance, as can be inferred from Holt and Laury (2002, 2005). They convincingly demonstrate that it may make a difference whether choices are hypothetical or for real money: Contrary to their findings on real payoffs, they detect no significant stake effect for hypothetical payoffs, which could explain the absence of a clear stake effect in Etchart-Vincent (2004) who investigates probability weights under hypothetical losses.

The paper is structured as follows. Section 2 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 3. The results of the estimation procedure are discussed in Section 4. Section 5 concludes the paper.

2 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 campusses of Peking University and Tsinhua University. Interested people had to register by email for one of two sessions conducted on the same day. Participants 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[1]. Twenty-eight lotteries offered low-stake outcomes ranging from 4 to 55 Chinese Yuan (CHN), another 28 lotteries entailed high-stake outcomes from 65 to 950 CHN[2]. Average earnings per subject amounted to approximately 323 CHN, including a show up fee of 20 CHN. Monetary incentives were substantial given the participants' average monthly disposable income of about 700 CHN. Besides, the low-stake outcomes were quite salient by themselves, as the expected payoff over low-stake lotteries amounted to about 16 CHN, considerably more than the going hourly wage rate. 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 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

[1]Instructions are available upon request.

[2]At the time of the experiment one Chinese Yuan equalled about 0.12 U.S. Dollars.

lotteries is presented in Table 1.

Subjects were entitled to one random draw from their low-stake decisions and to 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.

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.

For each lottery, a decision sheet, such as presented in Figure 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 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 Figure 1, her certainty equivalent

Figure 1: Design of the Decision Sheet

Decision situation: 22						
	Option A	Your Choice:				Option B Guaranteed payoff amounting to:
1		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	Profit of CHN 20 with probability 25% and profit of CHN 0 with probability 75%	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

would amount to 13.5 CHN.

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 minutes to complete the experiment. At the end of the experiment, one of each subject's low-stake as well as high-stake choices were randomly selected for payment. Subjects were paid in private afterward.

3 Econometric Model

This section discusses the specification of the econometric model, which is based on several building blocks: first, specifying the basic decision model; second, allowing for potentially different behaviors under low and high stakes; third, specifying the error term; and finally, accounting for heterogeneity in behavior by a finite mixture regression approach. At the end of this section we also briefly discuss some of the issues typically encountered when estimating finite mixture regression models.

3.1 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 two robust empirical phenomena: nonlinear probability weighting and loss aversion (Starmer, 2000). Therefore, a flexible approach, such as proposed by 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

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

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

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

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

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

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, the existence of loss aversion can be inferred from the difference in the domain-specific curvatures. According to Tversky and Kahneman, loss aversion, in the sense that "losses loom larger than corresponding gains", is present if $v'(x) < v'(-x)$ for $x \geq 0$ (Tversky and Kahneman (1992), p. 303). This is the case if the estimated β is significantly larger than α.

A variety of functions for modeling probability weights $w(p)$ have been proposed in the literature (Quiggin, 1982; Tversky and Kahneman, 1992; Prelec, 1998). We use the two-parameter specification suggested by Goldstein and Einhorn (1987) as well as by Lattimore, Baker, and Witte (1992):

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

We favor this specification because it has proven to account well for individual heterogeneity (Wu, Zhang, and Gonzalez, 2004)[3] 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 the probability weighting function deviates from linear weighting. The larger the value of δ, the more elevated is the curve, ceteris paribus. 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.

3.2 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:

$$\omega = \omega_0 + \omega_{HIGH} * HIGH,$$

with ω_0 representing the respective low-stake parameter. This step adds another six additional 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 indeed significant, the present model would be mis-specified, as the power functional, used for estimation, cannot account for changing relative risk aversion.

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

In this case, an alternative specification of the value function that can account for changing relative risk aversion would be called for. In particular, if the valuation of 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.

3.3 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 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 lottery range, which has to be taken account of by the estimation procedure. Second, 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. Third, lotteries in the gain domain may be evaluated differently from the ones in the loss domain. Therefore, we allow for domain-specific variance in the error term. These considerations yield the form

$$\sigma_{ig} = \xi_i |x_{1g} - x_{2g}|$$

for the standard deviation of the error term distribution, where ξ_i denotes an individual domain-specific parameter. Note that the model allows to test for both individual-specific and domain-specific heteroskedasticity by either imposing the restriction $\xi_i = \xi$, or by forcing all the ξ_i to be equal in both decision domains. Both restrictions are rejected by their corresponding likelihood ratio tests with p-values close to zero. Therefore, we control for all three types of heteroskedasticity in the estimation procedure, which adds two more parameters per individual, i.e. in total 306 parameters, to the econometric model.

3.4 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, estimating a single set of parameters is inappropriate and may render misleading results. For this reason, we estimate a finite mixture model in order to account for heterogeneity. 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. Given our assumptions on the distribution of the error term, the density of type c for the i-th individual can be expressed as

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

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

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

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.

3.5 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, Laird, and Rubin, 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. 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 Average Normalized Entropy (El-Gamal and Grether, 1995), 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 allow the researcher 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. In the context of mixture models, classical criteria, such as the Akaike Information Criterion AIC or the Bayesian Information Criterion BIC, are not suited for this purpose (Celeux, 1996). Celeux proposes to use entropy criteria based on the posterior probabilities of group assignment τ_{ic} instead, which are shown to perform better than the classical criteria. For example, if entropy increases when the number of different types is raised from two to three, group assignment of the individuals is less reliable, and the model tends to overfit the data. Therefore, the model with two types is to be preferred.

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[4] are discussed in Bruhin, Fehr-Duda, and Epper (2007).

4 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 heterogeneous 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.

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

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

probability losses. Therefore, at first glance, average behavior is adequately described by a model such as CPT.

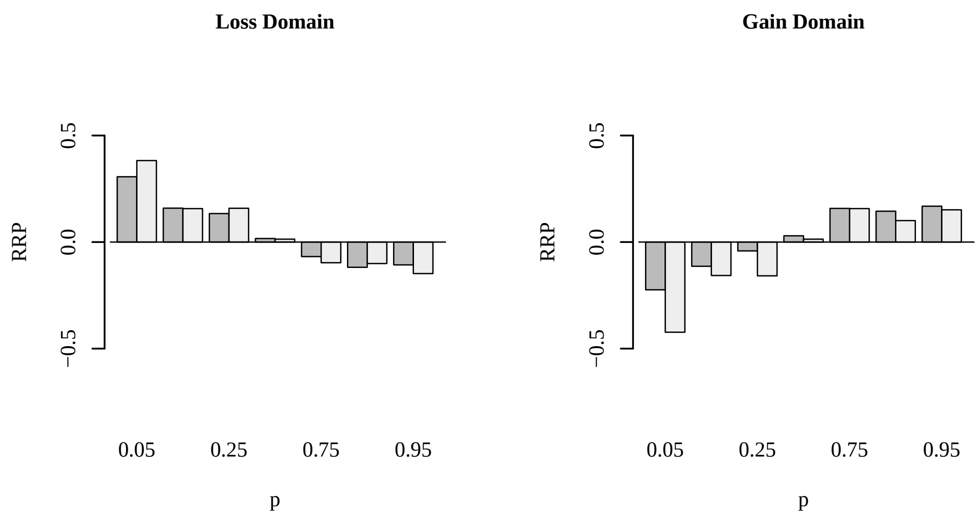
What the bar plots also reveal is that median relative risk premia differ substantially by stake level: When subjects' preferences exhibit increasing relative risk aversion, we should observe different low-stake and high-stake RRP, namely, high-stake choices should be relatively less risk tolerant than low-stake choices. Inspection of Figure 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, 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 conventional levels of significance: 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

Figure 2: Median Relative Risk Premia by Stake Size



Low stakes: light gray. High stakes: dark gray.

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 homogeneous preference model, the estimated curvatures of the value functions do not significantly change with rising stakes.

Support. Table 2 contains the parameter estimates for the decision model discussed in Section 3. For the time being, we focus on the average parameter estimates displayed in columns (1) and (4). The curvature parameters of the value functions over low stakes are denoted by α_0 for gains and β_0 for losses. 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 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: For homogeneous preferences, low-gain probability weights deviate more strongly from rational linear weighting than high-gain probability weights. No substantial change in probability weights is observed for losses.

Support. We first discuss the results 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 Figure 2. The *differences* in the observed stake-dependent RRP decrease markedly with increasing probability level, suggesting a substantial interaction effect[5].

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 1.307 to 0.979, induced by substantially less optimistic weighting of high-gain probabilities. Moreover, the change in the slope of the probability

[5]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 a kind of loss aversion in subjects' responses.

Table 2: Classification of Behavior: Parameter Estimates

	Gains				Losses		
	Pooled	EUT	Non-EUT		Pooled	EUT	Non-EUT
	(1)	(2)	(3)		(4)	(5)	(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					

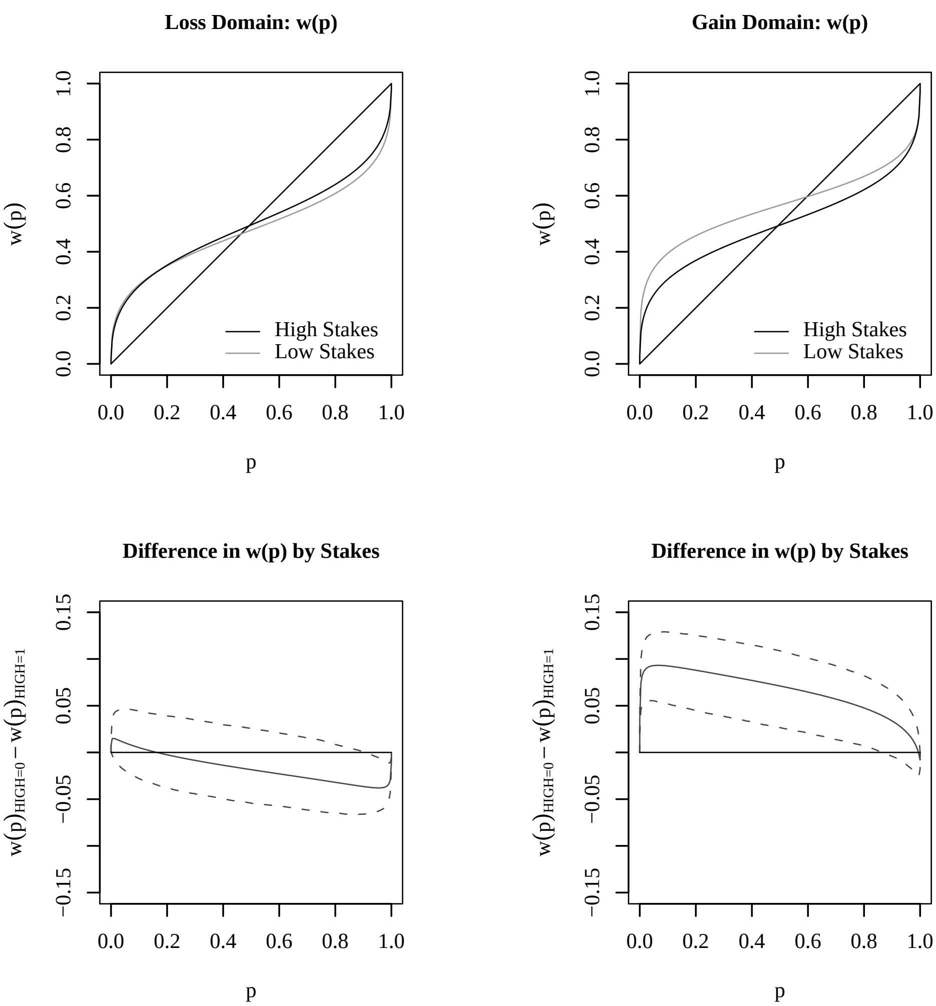
Standard errors in parentheses are based on the bootstrap with 2,000 replications.
Parameter vectors include estimates of $\hat{\xi}_i$ for domain- and individual-specific error variances.

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 Figure 3. The top panel of the figure shows, for each decision domain, the estimated probability weighting curves for low stakes, defined by $HIGH = 0$, plotted against the high-stake curves, defined by $HIGH = 1$. Evidently, the high-gain function is less elevated and slightly less strongly curved than the low-gain function, indicating a less pronounced departure from rational linear probability weighting for high gains.

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 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 Figure 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 right 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 Figure 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 Figure 3 shows that the 95%-condidence band for the difference in the stake-dependent probability weighting curves over losses includes the

Figure 3: Average Probability Weights by Stake Size



Dashed lines: 95%-confidence bands based on the percentile bootstrap method.

The index LS denotes $HIGH = 0$, HS denotes $HIGH = 1$.

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.

4.2 Heterogeneous Types of Behavior

So far 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. The first question to be answered concerns the number of different types present in the population. One way of dealing with this question is calculating a measure of entropy for varying numbers of groups and choosing the model with the lowest entropy.

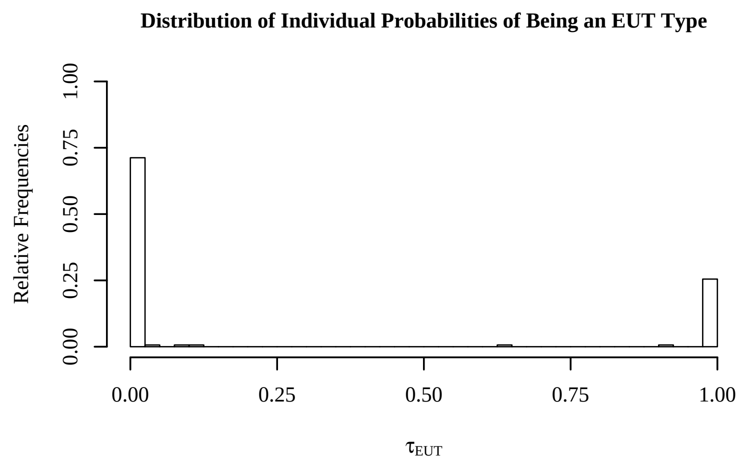
RESULT 4: There is substantive heterogeneity in individual risk preferences, which can be captured by two distinct types of behavior.

Support. The finite mixture regression model classifies individuals according to a given number of types. In order to evaluate the quality of classification, we calculated the Average Normalized Entropy ANE (El-Gamal and Grether, 1995) defined as

$$ANE = -\frac{1}{N}\sum_{i=1}^{N}\sum_{c=1}^{C}\tau_{ic}\log_C(\tau_{ic}),$$

for C groups and N individuals. Taking $\log_C$ normalizes the entropy measure to lie within $[0,1]$. If all the probabilities of individual group membership τ_{ic} are equal to zero or one, $ANE = 0$. In this case, all the individuals can be perfectly assigned to one group. $ANE = 1$ reflects maximum entropy, i.e. all the τ_{ic} are equal to $1/C$. Such a result indicates that group membership is totally ambiguous and that categorization has failed. For two groups,

Figure 4:



we find an ANE-value of 1.8% of the maximum entropy. Given this extremely low degree of ambiguity in our two-group classification, an improvement in entropy seems hardly possible. If the classification procedure worked better for three groups than for two groups, the average normalized entropy should be smaller for $C = 3$ than for $C = 2$. This is clearly not the case, as ANE amounts to 3.2% for the three-group classification[6]. So we can safely conclude that two types of behavior are sufficient to capture the essential characteristics of individual heterogeneity in risk taking.

The low value of ANE 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 Figure 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

[6]Moreover, it can be shown that, when three groups are allowed for, the minority group remains stable, and the majority one is divided up into two separate groups exhibiting qualitatively similar parameter estimates.

of close to 75% 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?

4.2.1 Minority Behavior

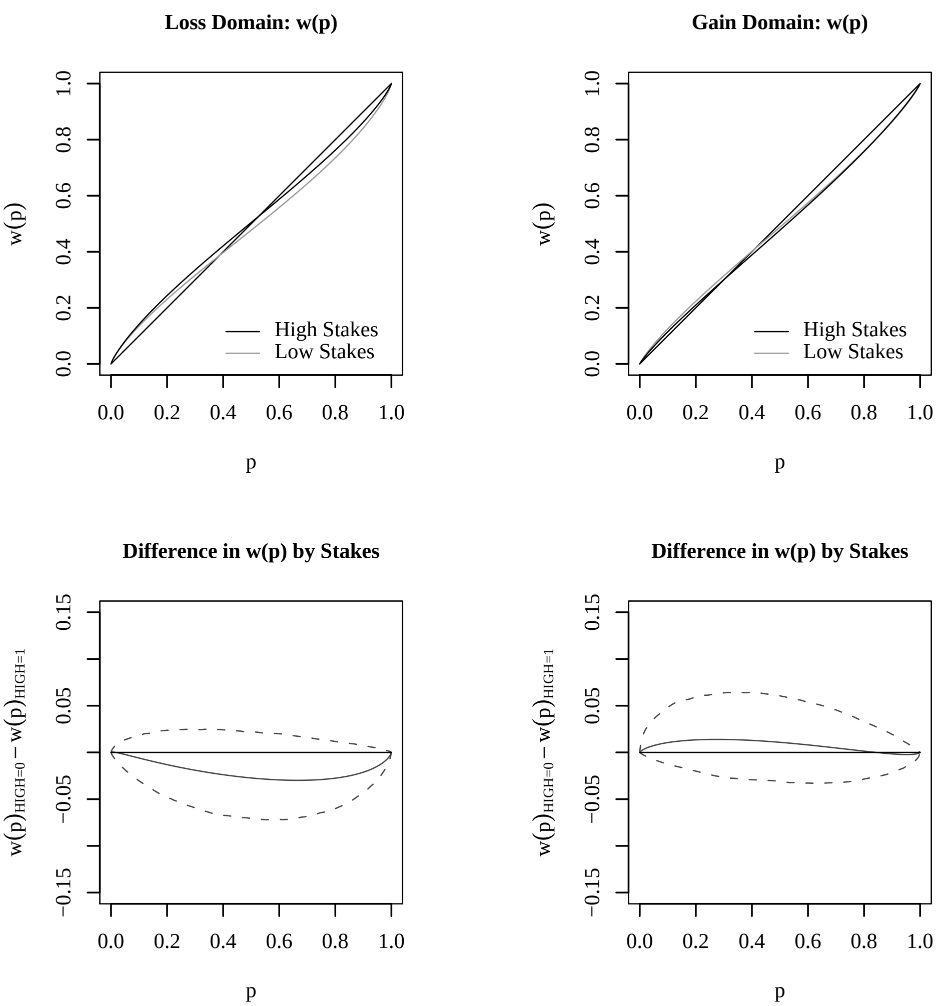
RESULT 5: The minority type, constituting about 27% of the subjects, can essentially be characterized as expected value maximizers over both low- and high-outcome ranges.

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 Figure 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 Figure 5 confirms that departures from linear probability weighting are insubstantial. Furthermore, for both gains and losses, no stake-size effect is visible in slope nor 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 Figure 5. These findings suggest that the minority type of decision makers behaves essentially as expected value maximizers, and therefore consistently with EUT. These conclusions, based on the estimation results, also bear out at the level of

Figure 5: Probability Weights by Stake Size: EUT Types



Dashed lines: 95%-confidence bands based on the percentile bootstrap method.

The index LS denotes $HIGH = 0$, HS denotes $HIGH = 1$.

observed behavior. The EUT types' median relative risk premia in the bottom panel of Figure 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.

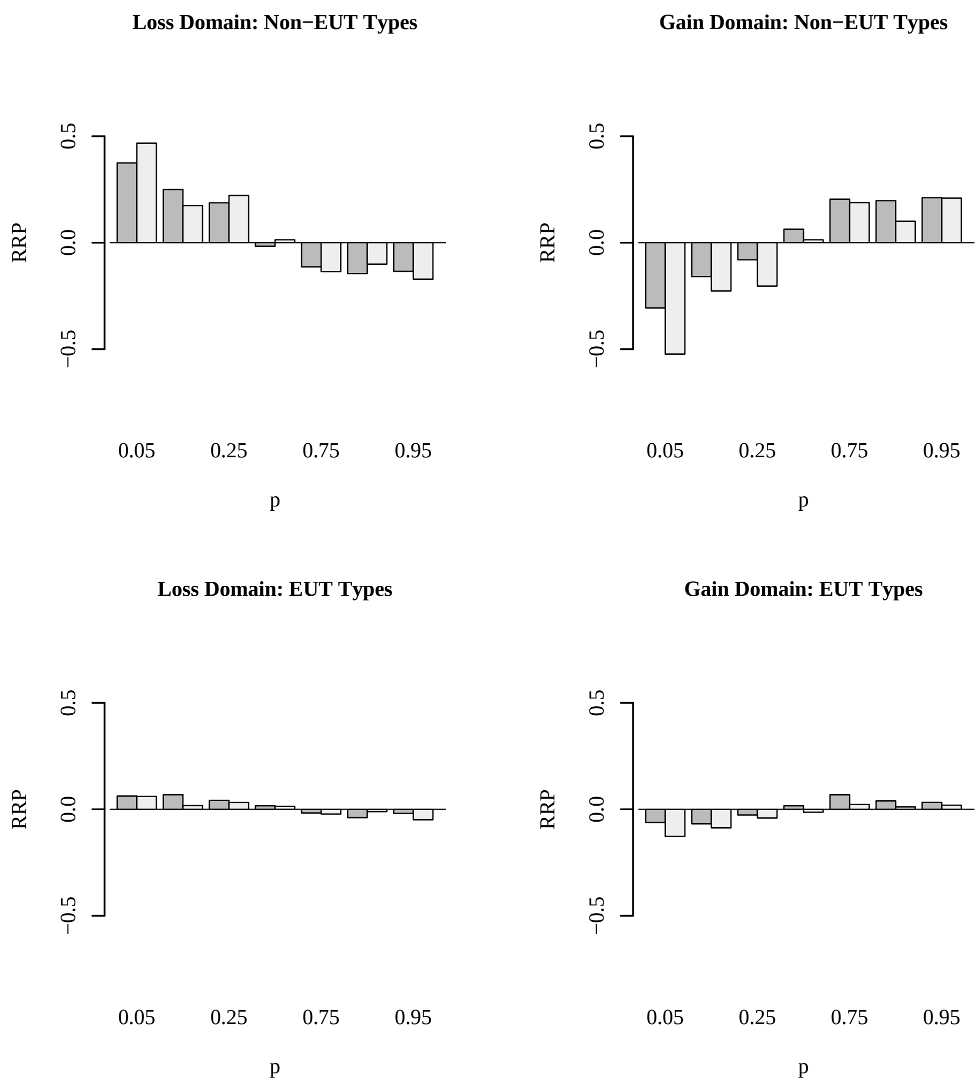
4.2.2 Majority Behavior

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 a 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 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.

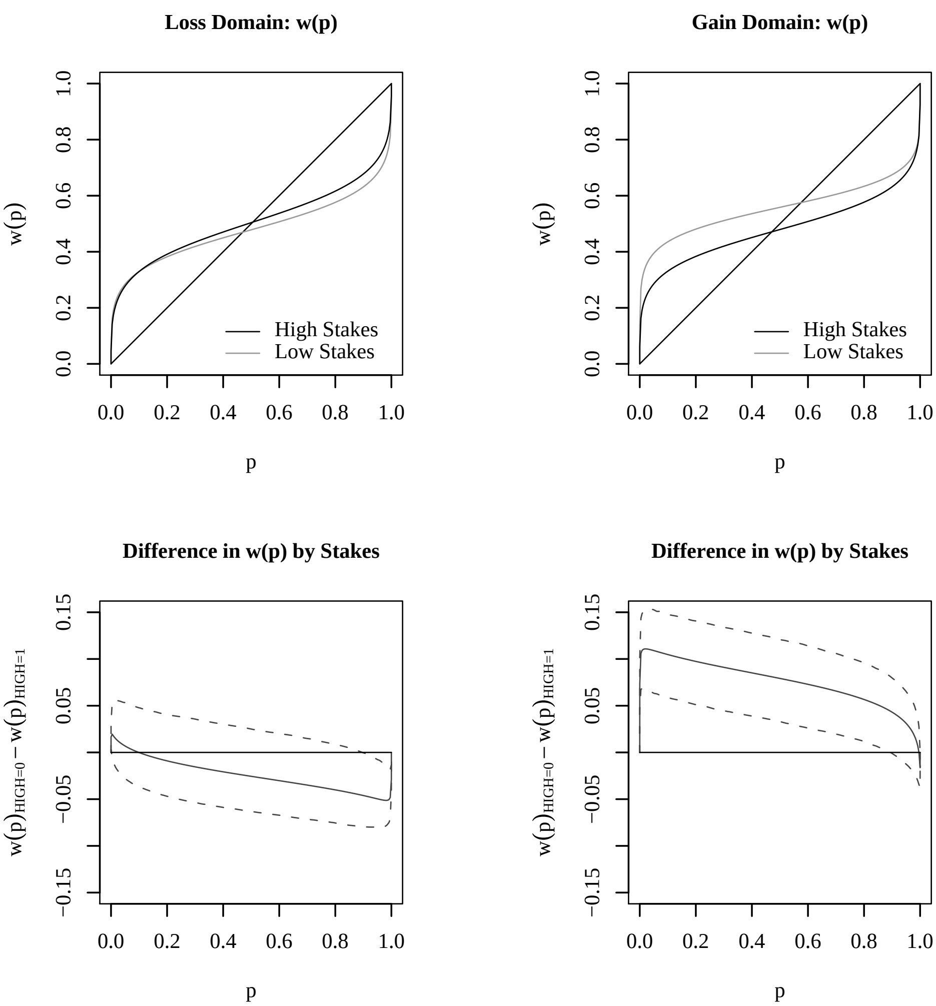
In order to examine the stake-specific probability weighting curves for the majority group we inspect the top panel of Figure 7. Our findings on the majority group's curves reflect the same patterns of stake sensitivity as the pooled ones do: For both gains and losses the curves are inverted S-shaped, 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 Figure 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 in the corresponding estimate for δ_{HIGH}, which amounts to -0.344 (see column (3) in Table 2). The high-gain probability weighting function is also slightly less curved than the low-gain one, as γ_{HIGH} is estimated to be 0.058.

Figure 6: Type-Specific Median Relative Risk Premia by Stake Size



Low stakes: light gray. High stakes: dark gray.

Figure 7: Probability Weights by Stake Size: Non-EUT Types



Dashed lines: 95%-confidence bands based on the percentile bootstrap method.

The index LS denotes $HIGH = 0$, HS denotes $HIGH = 1$.

The joint impact of these parameter changes is statistically significant, as the right hand side of the bottom panel of Figure 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 Figure 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 Figure 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.

5 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. The results of this investigation can be summarized as follows: 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. Since 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.

Contrary to previous attempts at explaining the increase in relative risk aversion over gains by changing attitudes toward monetary payoffs, this increase can be mainly attributed to a move of the average probability weighting function towards rational linear weighting. As the finite mixture regression analysis shows, this average effect is brought about by the behavior of a majority of decision makers who tend to weight probabilities of low-stake gains considerably more optimistically than probabilities of high-stake gains. Whereas these Non-EUT types' behavior is sensitive to payoff levels, the minority group's behavior, which is shown to be largely consistent with expected value maximization, is not.

These 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 payoffs in a material way. Our results suggest that this interaction effect is significant and substantial in the gain domain, which renders rank-dependent models, such as prospect theory, questionable when risk preferences over a wide range of outcomes are concerned.

In the field of applied economics, one of the most important issues concerns the substantive heterogeneity found in the population. This study has demonstrated that there are two distinct behavioral types who either weight probabilities near linearly or nonlinearly. Two clearly segregated groups of comparable size and characteristics were also found in two independent Swiss data sets (Bruhin, Fehr-Duda, and Epper, 2007) and, for choices over gains only, in a British data set (Conte, Hey, and Moffat, 2007), which suggests that this mix of preference types seems to be quite robust across times and cultures. This substantive kind of heterogeneity has to be taken into account when predicting behavior, as average parameter estimates may be quite misleading. Moreover, as the literature on the role of bounded rationality under strategic complementarity and substitutability has shown, the mix of rational and irrational actors may be crucial for aggregate outcomes (Haltiwanger and Waldman, 1985, 1989; Fehr and Tyran, 2005). This literature demonstrates that, depending on the nature of strategic interdependence, even a minority of players of a particular type may be decisive.

The researcher will, therefore, have to determine which one of the distinct behavioral groups identified in the population will most likely dominate aggregate outcomes. If she can safely conclude that the minority EUT type will dominate, stake dependence of risk preferences is not an issue, and expected value maximization may be the adequate model of decision making. If, however, she regards the Non-EUT types as decisive for aggregate outcomes, stake dependence might be a serious problem when gains are concerned. If actual stakes are much larger than the ones used for model estimation, predicting behavior on the basis of estimated model parameters might lead to a significant overestimation of risk tolerance. In order to get a handle on choices under substantial stakes, research will probably have to turn to field data to generate meaningful parameter estimates. Since our results suggest that stake-sensitivity is largely due to a change in probability weights, using more flexible specifications of utility functions, as proposed by Holt and Laury (2002), cannot adequately solve the problem of modelling risk preferences.

The issue of stake sensitivity is an important one for choices over gains. When evaluating potential losses, however, subjects seem to use different heuristics and decision rules than when evaluating gains, which renders them rather insensitive to stake size. This stability of behavior can be seen in both the EUT and the Non-EUT groups. While stake dependence is not an issue here, the researcher will still have to worry about heterogeneity and the ensuing type-dominance question. As Mason, Shogren, Settle, and List (2005) have pointed out, the common assumption of linear probability weights may lead to problematic policy recommendations, if nonlinear probability weighting is the dominant pattern of behavior. The results of this paper cast doubt on research strategies that do not take framing effects, nonlinear probability weighting and the substantive heterogeneity of decision makers into account.

Acknowledgments: We are particularly thankful to Mei Wang, Xiaofei Xie, Marc Schuerer, Ernst Fehr, George Loewenstein, and Roberto Weber, as well as to the participants of the ESA World Conference 2007 and the Behavioral Decision Research Seminar at Carnegie Mellon University.

References

BINSWANGER, H. P. (1981): "Attitudes toward Risk: Theoretical Implications of an Experiment in Rural India," *Economic Journal*, 91, 867–890.

BOSCH-DOMENECH, A., AND J. SILVESTRE (1999): "Does Risk Aversion or Attraction depend on Income? An Experiment," *Economic Letters*, 65, 265–273.

BRUHIN, A., H. FEHR-DUDA, AND T. F. EPPER (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*. W. Edwards (ed.), Utility: Measurement, Theory and Applications. Amsterdam, Kluwer, Nijhoff.

CELEUX, G. (1996): "An Entropy Criterion for Assessing the Number of Clusters in a Mixture Model," *Journal of Classification*, 13, 195–212.

CONTE, A., J. D. HEY, AND P. G. MOFFAT (2007): "Mixture Models of Choice Under Risk," *Department of Economics and Related Studies, University of York, Discussion Paper*, 2007/06.

DEMPSTER, A., N. LAIRD, AND D. RUBIN (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., AND D. M. GRETHER (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., AND J.-R. TYRAN (2005): "Individual Irrationality and Aggregate Outcomes," *Journal of Economic Perspectives*, 19(4).

Goldstein, W., and H. Einhorn (1987): "Expression Theory and the Preference Reversal Phenomena," *Psychological Review*, 94, 236–254.

Haltiwanger, J. C., and M. Waldman (1985): "Rational Expectations and the Limits of Rationality: An Analysis of Heterogeneity," *American Economic Review*, 75(3), 326–340.

——— (1989): "Limited Rationality and Strategic Complements: The Implications for Macroeconomics," *Quarterly Journal of Economics*, 104(3), 463–483.

Hey, J. D., and C. Orme (1994): "Investigating Generalizations of Expected Utility Theory Using Experimental Data," *Econometrica*, 62(6), 1291–1326.

Hogarth, R. M., and H. J. Einhorn (1990): "Venture Theory: A Model of Decision Weights," *Management Science*, 36(7), 780–803.

Holt, C. A., and S. K. Laury (2002): "Risk Aversion and Incentive Effects," *American Economic Review*, 92(5), 1644–1655.

——— (2005): "Risk Aversion and Incentive Effects: New Data without Order Effects," *American Economic Review*, 95(3), 902–904.

Houser, D., M. Keane, and K. McCabe (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., and M. Shehata (1992): "Examining Risk Preferences Under High Monetary Incentives: Experimental Evidence from the People's Republic of China," *American Economic Review*, 82(5), 1120–1141.

Kuehberger, A., M. Schulte-Mecklenbeck, and J. Perner (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., J. R. Baker, and A. D. Witte (1992): "The Influence of Probability on Risky Choice," *Journal of Economic Behavior and Organization*, 17, 377–400.

Markowitz, H. (1952): "The Utility of Wealth," *The Journal of Political Economy*, 60(2), 151–158.

Mason, C. F., J. F. Shogren, C. Settle, and J. A. List (2005): "Investigating Risky choices Over Losses Using Experimental Data," *Journal of Risk and Uncertainty*, 31(2), 187–215.

McLachlan, G., and D. Peel (2000): *Finite Mixture Models*. Wiley Series in Probabilities and Statistics.

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.

Saha, A. (1993): "Expo-Power Utility: A 'Flexible' Form for Absolute and Relative Risk Aversion," *American Journal of Agricultural Economics*, 75, 905–913.

Stahl, D. O., and P. W. Wilson (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., and D. Kahneman (1992): "Advances in Prospect Theory: Cumulative Representation of Uncertainty," *Journal of Risk and Uncertainty*, 5, 297–323.

Weber, B. J., and G. B. Chapman (2005): "Playing for Peanuts: Why is Risk Seeking More Common for Low-Stakes Gambles?," *Organizational Behavior and Human Decision Processes*, 97, 31–46.

Wu, G., J. Zhang, and R. Gonzalez (2004): *Decision under Risk*. D. Koehler and N. Harvey (Eds.), The Blackwell Handbook of Judgment and Decision Making, Oxford University Press.