

Risk attitude elicitation using a multi-lottery choice task: Real vs. hypothetical incentives

Iván Barreda-Tarrazona*

LEE and Economics Department, University Jaume I of Castellón (Spain)

Ainhoa Jaramillo-Gutiérrez

LINEEX and Applied Economics Department, University of Valencia (Spain)

Daniel Navarro-Martínez

LEE and Departments of Economics and Psychology, University of Warwick (UK)

Gerardo Sabater-Grande

LEE and Economics Department, University Jaume I of Castellón (Spain)

* Corresponding author: Iván Barreda Tarrazona. Tel.: +34964387633; fax: +34964387653. E-mail address: ivan.barreda@eco.uji.es. The article has benefited greatly from comments by Nikolaos Georgantzís, Cristiana Cerqueira and participants in the Valladolid Behavioral Finance Workshop, 2010. The experiments were made possible thanks to financial help from the Fundació Caixa Castelló - Bancaixa (P1 A2007-06). I. Barreda and G. Sabater acknowledge financial support by the Spanish Ministry of Science and Innovation: ECO2008-04636/ECON and I. Barreda also acknowledges the Spanish Ministry of Science and Innovation: ECO2010-18567 and the José Castillejo Grant for visiting ESRI Dublin in 2009.

Abstract

We present a bi-dimensional multi lottery choice task which can be used in order to elicit the agents' risk attitudes in financial environments. This task is implemented both with hypothetical and real monetary incentives in a between-subjects and a within-subjects experiment. We observe choices involving significantly lower risk aversion on aggregate when incentives are real. The differences grow with the stakes at play. We also obtain significant differences between hypothetical and real rewards in both utility weighting and probability weighting estimated parameters. We find that the use of hypothetical incentives in multi-lottery choice tasks for evaluating individual risk aversion can be misleading.

Resumen

Presentamos una lotería múltiple bidimensional que puede utilizarse para elicitación la actitud frente al riesgo en entornos financieros. Esta tarea fue implementada en dos tratamientos: entre e intra-sujetos, tanto bajo incentivos monetarios hipotéticos como reales. Observamos que los sujetos son significativamente menos adversos al riesgo cuando los incentivos que se les ofrecen son reales. Estas diferencias aumentan con la cuantía de los pagos en juego. También obtenemos diferencias significativas entre incentivos hipotéticos y reales en los parámetros estimados de ponderación de la utilidad y de ponderación de la probabilidad. El uso de incentivos hipotéticos para la evaluación de los niveles individuales de aversión al riesgo puede llevar a resultados engañosos.

Keywords: experimental economics, hypothetical bias, multi-lottery choice task.

JEL classification: C91, D81.

1- INTRODUCTION

In a recent survey, Harrison and Rutström (2008) affirm that reliable laboratory methods exist to determine the individual risk aversion of a subject and that these methods could be systematically employed to ensure greater control over tests and applications of theory that depend on risk attitudes. They clearly advocate in favor of saliently motivating subjects' responses. We want to investigate at the individual level the consequences of not doing it. A broadly used test among psychologists is Zuckerman's (1978) Sensation Seeking Scale (SSS), while economists mainly use the Holt and Laury (2002) type of binary lotteries (HL). The SSS asks about different types of risks, including financial risks, while HL is exclusively framed in the monetary domain. Both tests present the problem of uni-dimensionality of the risk aversion characterization of an individual. The Sabater-Grande and Georgantzís (2002) test (SGG) that we use allows us to obtain two parameters of the utility function in an efficient way.

The role of incentives in the context of individual decision making under risk and uncertainty has been recurrently explored in the literature. Since Edwards (1953) found, as we do, an increase in the willingness to take risks when participants play for real money, there have probably been more studies comparing "hypothetical" with "real" decisions in this context than in any other area of experimental economics. However, the issue is still far from settled and many articles are published still today using either method. Our aim in this study is to analyze the existence, direction, and practical relevance of the difference between risk aversion levels inferred under hypothetical and real incentives.

The general consensus among psychologists seems to be that hypothetical risky choices give a reasonable, qualitatively correct picture of real choices. Wärneryd (1996) supports their use in survey contexts. Wiseman and Levin (1996) carry out three experiments in which subjects make risky decisions under conditions of hypothetical or real consequences, finding no significant differences in any of them. Beattie and Loomes (1997) suggest that in simple pairwise choices, incentives appear to make very little difference with regard to performance. Also many economists, maybe influenced by the psychologists' experimental tradition as suggested by Harrison and Rutström (2008), do not always motivate the subjects monetarily when asking about their risk preferences. For instance, Kuhberger *et al.* (2002) find that the change from small incentives (hypothetical payoffs, real low payoffs) to high incentives (real high payoffs) leads to a difference in choices, but on the other hand, the same choices are made with real high payoffs than with hypothetical high payoffs. Dohmen *et al.* (2005) find that the answers to a general risk attitude question predict actual behavior in a lottery quite well. Also Faff *et al.* (2008) find no significant differences between using hypothetical or real payoffs when comparing financial risk tolerance with risk aversion.

However, the standard experimental economics methodology (Smith, 1982) advocates for salient economic rewards when designing an experiment and many studies report different results with hypothetical and real incentives. It is assumed that if subjects do not consider hypothetical gains seriously, they may be tempted to take more risks (or be

less risk averse) than when they are really likely to win. Camerer and Hogarth (1999) review 74 studies comparing behavior of experimental subjects who were not paid, or were paid low or high financial incentives according to their performance. They conclude, contrary to us, that when incentives are low, subjects declare that they would be more risk-loving than they actually are when incentives are increased. Also Etchart-Vincent and l'Haridon (2008) find that subjects exhibit more risk seeking when choices are hypothetical than real.

Holt and Laury (2002, 2005, 2008) find that increasing the size of real payoffs leads subjects to behave in a more risk averse manner both in the gain and the loss domain, while with hypothetical payments, more than half of the subjects who are risk averse for gains turn out to be risk seeking for losses.

Our results are in line with the studies in the literature which claim a difference between hypothetical and real payments. However, in contrast with most previous studies, we observe choices which are, on average, less risk averse when payments are real.

A within-subject design is more reliable than a between-subject design but it presents the potential bias of a carryover effect across sessions, which is very difficult to control even taking into account order effects, since, once the subject has been incentivized to think seriously about his risk preference, he will probably remember and try to be consistent with his decision even if asked again hypothetically. So we opted for using a between subjects and a within subjects design and cross-check in this way for the robustness of our results, with the advantage of having relatively many data available under both conditions from the experiments that we have carried out. In fact, a total of 786 subjects participated in our lotteries and 402 of them received real rewards for their decisions. No other study comparing hypothetical with real incentives in risk aversion elicitation has a comparable sample size.

Our results clearly advocate in favor of saliently motivating the answers of the risk-aversion test and the elicited level of risk aversion significantly decreases with respect to the case of no payment.

In the next section we explain in detail the experimental design. Then, in Section 3, we present the results. Conclusions and references follow.

2- EXPERIMENTAL DESIGN AND ECONOMETRIC MODEL

We organized two treatments. In the between subjects treatment (BST) our subjects were a relatively large sample of 695 subjects who were volunteers recruited among the undergraduate students of Business Administration from University Jaume I in Spain. From them, 384 subjects received no money for the lottery decision task and 311 subjects faced the real monetary consequences of the lottery that they had chosen.

In the within subjects treatment (WST) 91 Business Administration students also voluntarily recruited from the same university, who did not participate in the previous treatment, were presented the same lottery decisions as in the BST but they had to face both conditions: first by taking hypothetical decisions and, about one year later, repeating the test under real payment for the lotteries. The temporal stability of

estimates of risk aversion has been studied in detail by Harrison *et al.* (2005) and Andersen *et al.* (2008). Their results show evidence of stability for appropriately built risk aversion measures for periods up to one and a half year, if the personal socioeconomic conditions of the subjects are not importantly changed. Baucells and Villasís (2010) find some evidence of individual changes in a three month period, but they did not pay their subjects, which as we show in the present article, can lead per se to inconsistent decisions.

Apart from comparing with the between-subjects design in order to check for consistency, in the within-subjects treatment we also introduced a long time span in order to minimize possible carryover effects in the latter treatment. Average earnings in the case of real payments were 6€, the lotteries were explained and completed in about 10 minutes.

The experiments in which the lotteries were played involved no show up fee and no randomized payment. Therefore, our results are not very dependent on the possible sample selection biases pointed out by Harrison *et al.* (2009), particularly for the within subjects treatment, where the distribution of subjects' risk aversion levels is exactly the same.

Rather than usual tests based on binary choice tasks à la Holt and Laury, subjects were presented with the multi-lottery choice task, SGG, which is more appropriate for our purposes, due to the variety of results it produces¹. The task is designed to capture efficiently two dimensions of a subject's preferences towards risky choice. i) First, it distinguishes between risk neutral or loving subjects and subjects with different degrees of risk aversion as other lotteries do. ii) Second, the test explores the subjects' reaction to an increase in the magnitude of the risk compensation, that is, an increase in the stakes at play. In fact, by asking our subjects to take four decisions, we get four points of their utility function depending on the size of the compensation for risk, while the most widely used method gets one point after having asked multiple, normally ten, choices.

As shown in Table 1, the SGG task involves four panels of ten lotteries each. Each lottery $j = 1, \dots, 10$ entails a chance p_j of earning X € (else nothing). Each participant in our experiment had to choose one of the ten lotteries for each of the four panels, presented simultaneously to them.

After choices were collected, a four-sided die determined the panel which would be paid in the case of real payments. Subjects choosing the certain payoff in the selected panel were paid 1€. Subsequently, a 10-sided die was thrown to determine the "winning-lottery threshold". If the result of casting the die was 0, no payment was made to those having chosen a probabilistic payoff, if the result was any other number between 1 and 9, those subjects having chosen a loss probability lower or equal than that number divided by 10 got the prize corresponding to the probability chosen, the others got 0. Each one of the 4 panels is constructed using a certain payoff, $c = 1$ €, and the expected earnings, $p_j X_j$, are increased by a ratio t times the probability of not winning, $1 - p_j$, as implied by the formula: $p_j X_j = c + t(1 - p_j)$. That is, an increase in the probability of the unfavorable outcome is linearly compensated by an increase in the expected payoff.

¹ See García-Gallego *et al.* (2011) for details.

We use four different risk premium parameters in the four panels, $t = 0.1, 1, 5, \text{ and } 10$, implying an increase in the return of risky choices as we move from one panel to the next.

In order to study the subjects' reaction to an increase in t , that is, the magnitude of the risk compensation, we define the elasticity of the probability chosen in panel $i = 2, 3, \text{ and } 4$ to the increase in the risk premium as:

$$e_{p,t}^i = \frac{\frac{p_i - p_{i-1}}{p_{i-1}}}{\frac{t_i - t_{i-1}}{t_{i-1}}} \quad [1]$$

Additionally we define $e_{p,t}^{max} = \frac{\frac{p_4 - p_1}{p_1}}{\frac{t_4 - t_1}{t_1}}$ as the elasticity of the probability chosen in panel 1 to the maximum increase in risk compensation, occurring in panel 4².

Assuming, for instance, a constant relative risk aversion (CRRA) utility function, $U(X_j) = \frac{X_j^{(1-\alpha)}}{1-\alpha}$, where $j=1, 2, \dots, 10$, it can be checked that a subject maximizing the expected utility $EU(p_j X_j) = \frac{p_j^\alpha (c+t(1-p_j))^{1-\alpha}}{1-\alpha}$ would choose the lottery j with the probability closest to $\hat{p} = \alpha(1 + c/t)$.

On the one hand this confirms the intuitively expected outcome that the lower the probability of winning that the subject chooses, the less risk averse he is, whereas risk neutral/loving subjects would choose $p_j = 0.1$ in all panels. On the other hand, it predicts that the subject should choose riskier lotteries as we move from panel 1 to panel 4. Thus, for risk-averse expected utility maximizing subjects, their sensitivity to the attraction implied by a higher risk compensation t can be approximated by the difference in their choices across subsequent panels.

Our multi-lottery approach also allows us to estimate maximum likelihood models of utility functions in a similar way to Harrison and Ruström (2009). However, we have to adapt a structural model of binary choice to more than two categories, given that in SGG test we have ten possible choices.

First, we estimate a CRRA utility function using SGG lotteries data and assuming expected utility theory (EUT). We assume that utility for a subject n is defined by:

$$U_n(X_j) = \frac{X_j^{(1-\alpha_n)}}{1-\alpha_n} + \varepsilon_{nj} \quad [2]$$

Where, X_j is the prize of lottery j , α_n is the utility weighting parameter and ε_{nj} is the stochastic error, with expected value $E(\varepsilon_{nj}) = 0 \forall n, j$.

Under EUT, the value associated with X_j satisfies:

² The elasticities obtained from our data are shown in Table 3 and analyzed in the Results section.

$$EU_{nj} = EU_n(p_j X_j) = p_j \frac{X_j^{(1-\alpha_n)}}{1-\alpha_n} \quad [3]$$

The probability of a subject n selecting lottery j over all other possible lotteries is:

$$p(j) = p(EU_{nj} > EU_{nk}) \forall k \neq j \quad [4]$$

Assuming that ε_{nj} follows an independently and identically distributed (IID) logistic distribution:

$$p(j) = \frac{e^{EU_{nj}}}{\sum_{k=1}^{10} e^{EU_{nk}}} \quad [5]$$

For instance, the log likelihood of the multinomial logit model is:

$$\ln \mathcal{L} = \sum_{n=1}^N \sum_{j=1}^{10} z_{nj} \ln \left(\frac{e^{EU_{nj}}}{\sum_{k=1}^{10} e^{EU_{nk}}} \right) \quad [6]$$

Where $z_{nj} = 1$ if individual n chooses lottery j , $z_{nj} = 0$ otherwise.

Second, we estimate maximum likelihood models of utility functions assuming Rank Dependent Utility Theory (RDUT). We consider the Tversky and Kahneman (1992) probability weighting function:

$$w_n(p_j) = \frac{p_j^{\gamma_n}}{(p_j^{\gamma_n} + (1-p_j)^{\gamma_n})^{\frac{1}{\gamma_n}}} \quad [7]$$

where γ_n is the probability weighting parameter, that is, each subject can interpret the same probability in a personal way.

Under RDUT, the value associated with a lottery X_j satisfies:

$$EU_n(p_j X_j) = w_n(p_j) \cdot U_n(X_j) = \frac{p_j^{\gamma_n}}{(p_j^{\gamma_n} + (1-p_j)^{\gamma_n})^{\frac{1}{\gamma_n}}} \cdot \frac{X_j^{(1-\alpha_n)}}{1-\alpha_n} \quad [8]$$

We estimate again equation [6] using now equation [8], thus obtaining both the utility weighting parameter $\hat{\alpha}$ and the probability weighting parameter $\hat{\gamma}$.

3- RESULTS

In table 2 we present descriptive statistics of the choices made by panel, treatment and reward method (hypothetical, N=384; real money, N=311). Additionally, in Figure 1 and Figure 2 we present histograms of subjects' probability choices by panel and reward method corresponding to the between and the within-subjects treatments, respectively.

The median in the real reward panels is around 0.4 while it is 0.5 in the hypothetically paid panels and this difference is always significant using Mann Whitney tests in the BST and Wilcoxon tests in the WST. Specifically, we observe in Table 4 that with real rewards, the probability chosen by subjects is significantly lower than the probability chosen with hypothetical rewards in both treatments, with the exception of panel 4 in the WST.

Result 1: *We observe that in both the between and the within-subjects treatments our subjects choose on average riskier lotteries in the SGG test when given real payments as compared to hypothetical ones.*

Additionally, using a Levene test we find, as we can see in Table 5, that in the between-subjects treatment the variance of the probabilities chosen by subjects in any panel is significantly higher with hypothetical payments than with real ones. In contrast, in the within-subjects treatment, we obtain this finding for panel 4 only.

Result 2: *We find that in the SGG multiple lottery task real rewards generate more concentrated choices than hypothetical rewards.*

Using a Kolmogorov-Smirnov test (see Table 6) we obtain that hypothetical and real rewards generate significantly different distributions of observations³. From Figures 1 and 2 we can observe that with real payments the distribution generally shifts to the left, implying lower levels of risk aversion, and kurtosis grows, reflecting lower variance in the decisions.

Comparing the elasticities of choices with hypothetical and real payments using a Mann-Whitney test (see Table 4), we obtain that, in the BST, subjects' reaction to an increase in risk compensation is larger when rewards are real than hypothetical, with the exception of e4. In the WST this effect is confirmed only for $e_{p,t}^{\max}$.

Result 3: *In the BST the change in subjects' chosen probability from panel to panel is relatively greater when payments are real than when they are hypothetical. In the WST only the change between the first and the last panel is significant.*

Apart from calculating the elasticity we have estimated the multinomial logit models presented in Section 2, both under EUT and RDUT and for our two different treatments: real vs. hypothetical payment. We estimate them by maximum likelihood using the clustering method that allows for the possibility of correlation between responses by the same subject: the standard errors of the estimates are corrected for the possibility that the four responses are clustered for the same subject.

Estimation results are reported in Table 7. Under EUT, the results for the WST and BST are analogous. In the WST the average of the CRRA parameter estimate α is 0.600 with real payment and 0.638 with hypothetical payment. This difference is significant, confirming again our Result 1 that subjects are more risk averse when payments are hypothetical. The results are equivalent for the BST: α is 0.621 with real payment and 0.665 with hypothetical payment, and this difference is also significant. These values are in accordance to those obtained by Harrison *et al.* (2009).

³ With the exception of panels 3 and 4 in the WST.

Under RDUT, the CRRA coefficient α is again 0.600 with real payment and 0.633 under hypothetical payment in the WST. In the BST these values are 0.619 and 0.658 respectively. All these results are very similar to those under EUT. Regarding the estimates of the probability weighting parameter γ , we obtain a value of 0.647 with real payment and 0.678 with hypothetical payment for the WST. This difference is significant and indicates that the overweighting (underweighting) of small (large) probabilities is more pronounced under real payment. We can observe these effects in Figure 3.

In the BST we obtain equivalent results, estimated γ being 0.638 under real payment and 0.681 under hypothetical payment (see Figure 4).

Result 4: *Overweighting (underweighting) of small (large) probabilities is greater under real payment.*

There are no significant differences between WST and BST, neither in α nor in γ , showing the robustness of the result. To our knowledge, this is the first paper showing that probability weighting is affected depending on whether real or hypothetical rewards are used⁴.

4 – CONCLUSION

We have analyzed the existence, direction and practical relevance of the difference between risk aversion levels inferred under hypothetical and real incentives. Measuring individuals' risk aversion can prove very useful in order to interpret the decisions they take under financial risks. Different tests have been developed both in the psychological and in the economic literature to this aim. We present results based on Sabater-Grande and Georgantzís (2002) multi-lottery choice tests of risk attitude. In contrast to previous studies we obtain that when incentives are real subjects are less risk averse than when they are hypothetical.

Apart from explaining the characteristics of the test, we show that the way in which it is applied is also crucial. If incentives are hypothetical, the answers are noisier, less sensitive to changes in the stakes at play, and show a greater level of risk aversion than if subjects are monetarily motivated. The SGG test we use has good properties allowing us to efficiently obtain two parameters of the utility function of the agent using Rank Dependent Utility Theory. The estimated value of the utility weighting parameter α is significantly lower under real than under hypothetical payments. This means that our subjects are less risk-averse under real incentives and the estimated value for α (around 0.60 for real payment) is in line with the values obtained by Harrison *et al.* (2009) in different studies with other samples. We also obtain differences for the estimated value of γ , the probability weighting parameter in RDUT, this being also significantly lower under real payments. The estimated value of γ (close to 0.64 for real payment) implies the typical overweighting of the small probabilities and underweighting of large probabilities by our subjects.

⁴ Using a between subjects design, Harrison *et al.* (2010) do not find any significant hypothetical bias for purchasing managers assuming a rank dependent utility model.

We obtain these results from a sample of subjects larger than any other comparable study and we use a double design: both within and between-subjects treatments were implemented so that the smaller within-subjects treatment served as robustness check of the between-subjects treatment, obtaining similar results.

APPENDIX: TABLES AND FIGURES

<i>p</i>	<i>X €</i>	<i>p</i>	<i>X €</i>	<i>p</i>	<i>X €</i>	<i>p</i>	<i>X €</i>
1	1	1	1	1	1	1	1
0,9	1,12	0,9	1,2	0,9	1,67	0,9	2,20
0,8	1,27	0,8	1,50	0,8	2,50	0,8	3,80
0,7	1,47	0,7	1,90	0,7	3,57	0,7	5,70
0,6	1,73	0,6	2,30	0,6	5	0,6	8,30
0,5	2,10	0,5	3	0,5	7	0,5	12
0,4	2,65	0,4	4	0,4	10	0,4	17,50
0,3	3,57	0,3	5,70	0,3	15	0,3	26,70
0,2	5,40	0,2	9	0,2	25	0,2	45
0,1	10,90	0,1	19	0,1	55	0,1	100
PANEL 1		PANEL 2		PANEL 3		PANEL 4	

Table 1: Lottery probabilities (*p*) and payoffs (*X€*)

			Obs.	Mean	Median	Std. Dev.
BST	Hypothetical	Panel 1	384	0.5765625	0.6	0.2778557
		Panel 2	384	0.5171875	0.5	0.2387225
		Panel 3	384	0.5106771	0.5	0.2300314
		Panel 4	384	0.475	0.5	0.2504304
		Mean	384	0.5198568	0.525	0.205881
	Real	Panel 1	311	0.3710611	0.4	0.2143223
		Panel 2	311	0.3977492	0.4	0.1695195
		Panel 3	311	0.4241158	0.4	0.156258
		Panel 4	311	0.4041801	0.4	0.1596828
		Mean	311	0.3992765	0.4	0.1400
WST	Hypothetical	Panel 1	91	0.5043956	0.5	0.2699432
		Panel 2	91	0.4934066	0.5	0.1878547
		Panel 3	91	0.4857143	0.5	0.1524405
		Panel 4	91	0.443956	0.5	0.2061331
		Mean	91	0.4818681	0.475	0.1437139
	Real	Panel 1	91	0.378022	0.3	0.235609
		Panel 2	91	0.4065934	0.4	0.1678638
		Panel 3	91	0.4461539	0.4	0.1249957
		Panel 4	91	0.4175824	0.4	0.1487678
		Mean	91	0.4120879	0.4	0.1307756

Table 2: Descriptive statistics corresponding to probability choices by subjects

			Obs.	Mean	Std. Dev.
BST	Hypothetical	$e_{p,t}^2$	384	0.008006	0.082972
		$e_{p,t}^3$	384	0.051577	0.293277
		$e_{p,t}^4$	384	-0.018736	0.501755
		$e_{p,t}^{\max}$	384	0.000805	0.010033
	Real	$e_{p,t}^2$	311	0.355647	0.086257
		$e_{p,t}^3$	311	0.059369	0.184351
		$e_{p,t}^4$	311	0.050741	0.553799
		$e_{p,t}^{\max}$	311	0.004727	0.104512
WST	Hypothetical	$e_{p,t}^2$	91	0.259637	0.090199
		$e_{p,t}^3$	91	0.031768	0.114136
		$e_{p,t}^4$	91	-0.003048	1.040799
		$e_{p,t}^{\max}$	91	0.003562	0.013559
	Real	$e_{p,t}^2$	91	0.035246	0.078817
		$e_{p,t}^3$	91	0.059584	0.127365
		$e_{p,t}^4$	91	-0.015371	0.329843
		$e_{p,t}^{\max}$	91	0.005880	0.011435

Table 3: Descriptive statistics corresponding to elasticities of probabilities chosen to risk returns

Comparison HR-RR	Probability	P-value	Elasticity	P-value
BST	Panel 1	0.000 (+)	e_2	0.0000 (-)
	Panel P2	0.000 (+)	e_3	0.0003 (-)
	Panel P3	0.000 (+)	e_4	0.2861 (=)
	Panel P4	0.000 (+)	e_{\max}	0.0000 (-)
WST	Panel P1	0.001 (+)	e_2	0.1834 (=)
	Panel P2	0.003 (+)	e_3	0.1413 (=)
	Panel P3	0.0158 (+)	e_4	0.0933 (=)
	Panel P4	0.9730 (=)	e_{\max}	0.0319 (-)

Table 4: P-values corresponding to Mann Withney (BST) and Wilcoxon test (WST)

HR: hypothetical rewards; RR: real rewards; (+) indicates HR>RR; (-) indicates HR<RR; (=) indicates HR=RR

Comparison HR-RR	Panel	Probability	Elasticity
BST	P1	0.0000 (+)	0.0049 (-)
	P2	0.0000 (+)	0.3038 (=)
	P3	0.0000 (+)	0.0245 (-)
	P4	0.0000 (+)	0.0005 (-)
WST	P1	0.1498 (=)	0.8959 (=)
	P2	0.6881 (=)	0.3729 (=)
	P3	0.4984 (=)	0.1799 (=)
	P4	0.0039 (+)	0.5704 (=)

Table 5: P-values corresponding to Robust test for equality of variance (Levene test)

HR: hypothetical rewards; RR: real rewards; (+) indicates HR>RR; (-) indicates HR<RR; (=) indicates HR=RR

Comparison HR-RR	Panel	Probability	Elasticity
BST	P1	0.000 (≠)	0.000 (≠)
	P2	0.000 (≠)	0.000 (≠)
	P3	0.000 (≠)	0.033 (≠)
	P4	0.000 (≠)	0.000 (≠)
WST	P1	0.001 (≠)	0.345 (=)
	P2	0.010 (≠)	0.255 (=)
	P3	0.130 (=)	0.453 (=)
	P4	0.578 (=)	0.017 (≠)

Table 6: *P-values corresponding to Kolmogorov-Smirnov test*

HR: hypothetical rewards; RR: real rewards; (≠) indicates HR≠RR; (=) indicates HR=RR

WST		Expected utility		Rank-Dependent utility	
		Coefficient	Std. Errors	Coefficient	Std. Errors
α	<i>Real</i>	.6004834	.0100422***	.6002101	.0095474***
	<i>Hypothetical</i>	.6390580	.0108721***	.6341118	.0104128***
γ	<i>Real</i>			.6477575	.0096938***
	<i>Hypothetical</i>			.6784579	.0144159***
		$H_0: \alpha_{Real} = \alpha_{Hypothetical}$ <i>p-value= 0.003</i>		<i>p-value= 0.006</i>	
		$H_0: \gamma_{Real} = \gamma_{Hypothetical}$		<i>p-value= 0.028</i>	

BST		Expected utility		Rank-Dependent utility	
		Coefficient	Std. Errors	Coefficient	Std. Errors
α	<i>Real</i>	.6213423	.0049761***	.6194322	.0047398***
	<i>Hypothetical</i>	.6654572	.0048251***	.6585140	.0045997***
γ	<i>Real</i>			.6389014	.0053369***
	<i>Hypothetical</i>			.6813902	.0095887 ***
		$H_0: \alpha_{Real} = \alpha_{Hypothetical}$ <i>p-value= 0.000</i>		<i>p-value= 0.000</i>	
		$H_0: \gamma_{Real} = \gamma_{Hypothetical}$		<i>p-value=0.000</i>	

(***)significant at 1% confidence level

Table 7: *Parameters estimates of Expected utility and rank-dependent utility theories*

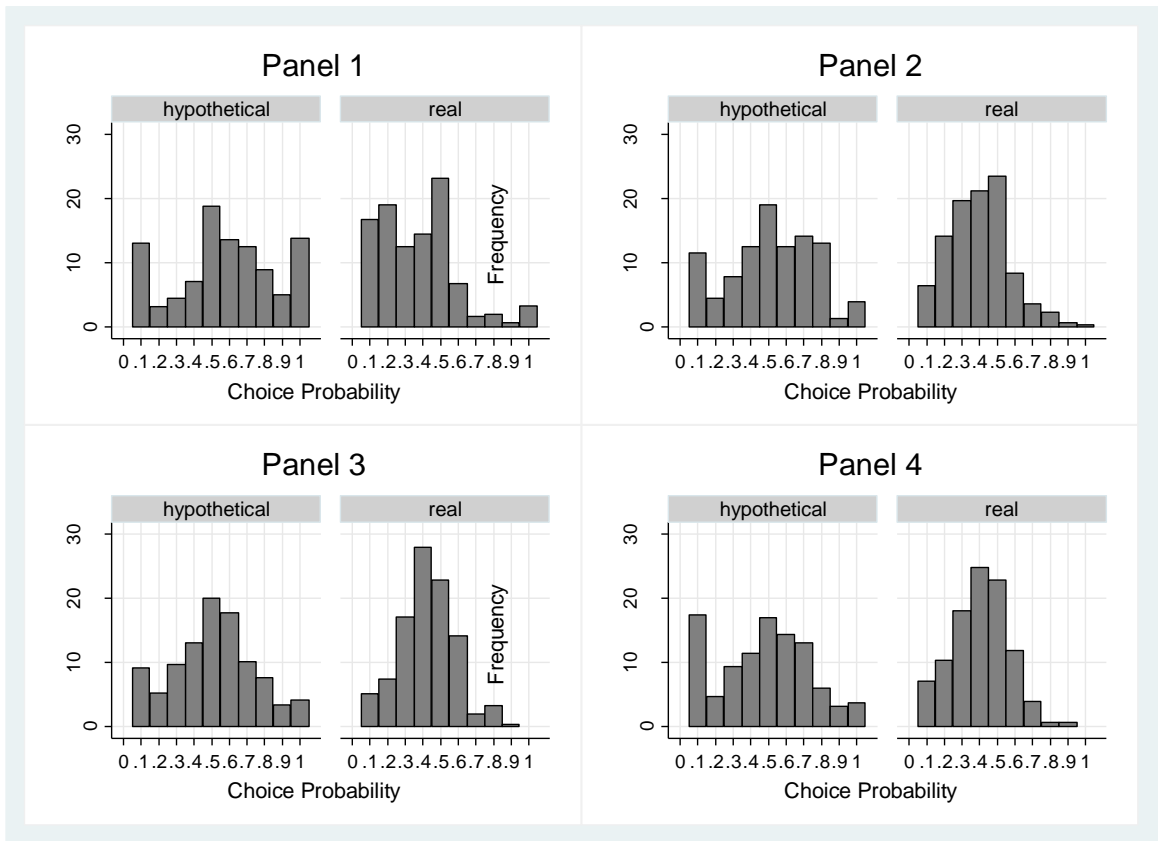


Figure 1: *Between subjects: Histograms of subjects' probability choices by panel.*

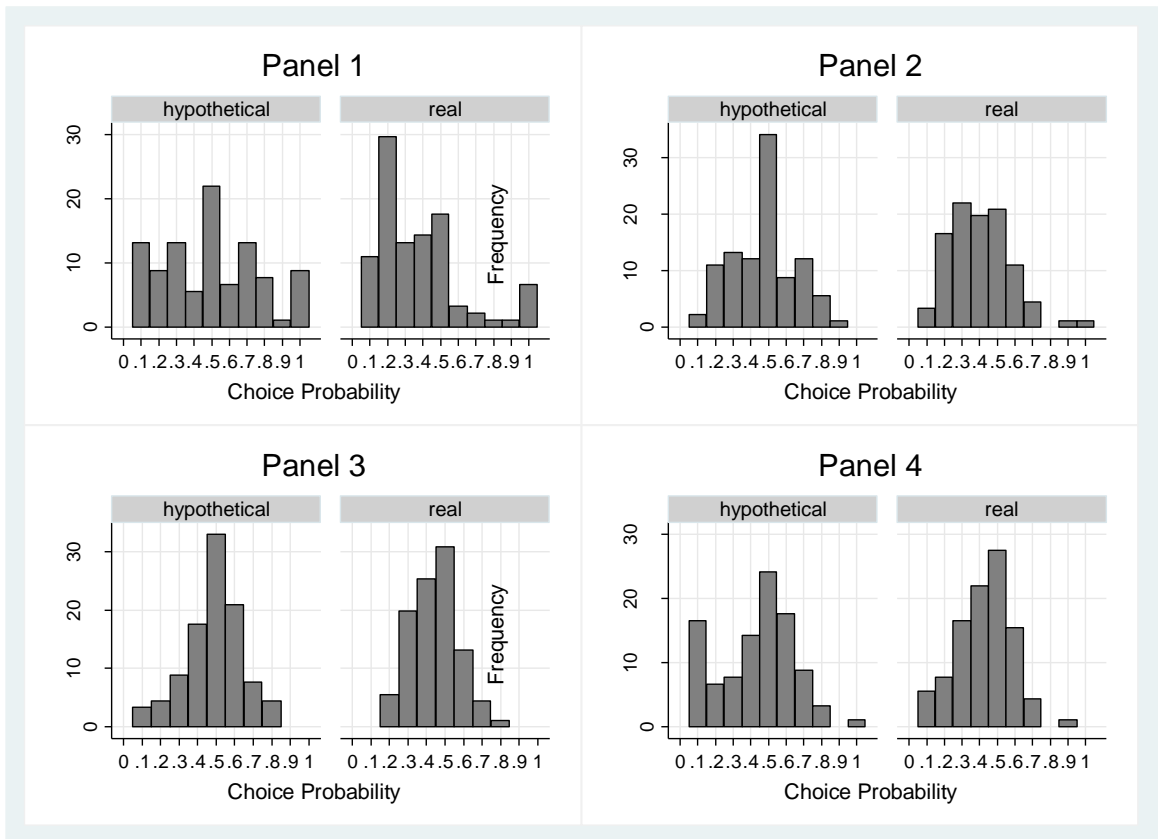


Figure 2: *Within subjects: Histograms of subjects' probability choices by panel.*

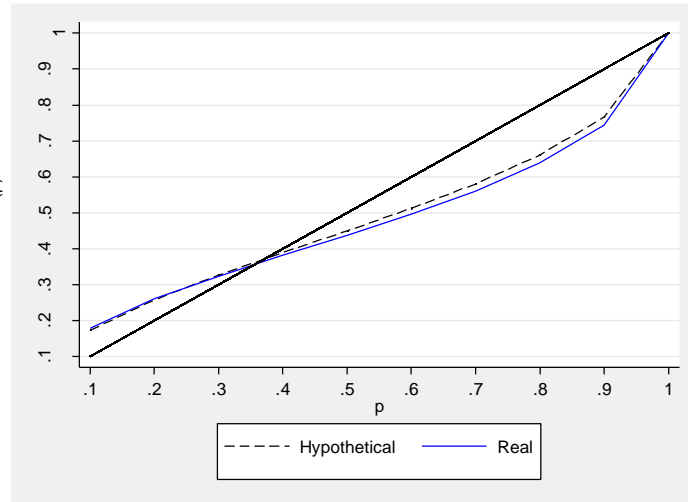


Figure 3: Within subjects estimated probability weighting functions with hypothetical and real payment.

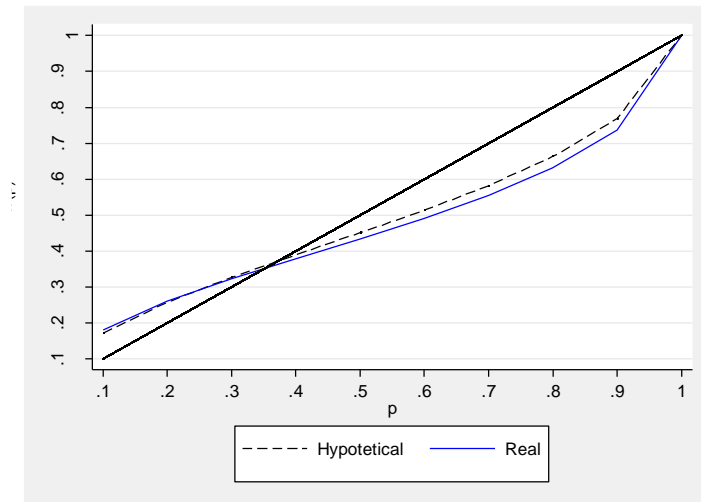


Figure 4: Between subjects estimated probability weighting functions with hypothetical and real payment.

REFERENCES

- Andersen, S., Harrison, G.W., Lau, M.I. and E.E. Rutström, 2008. Lost in state space. Are preferences stable? *International Economic Review* 49, 1091-1112.
- Baucells, M. and A. Villasís, 2010. Stability of risk preferences and the reflection effect of prospect theory, *Theory and Decision* 68, 193-211.
- Beattie, J. and G. Loomes, 1997. The impact of incentives upon risky choice experiments, *Journal of Risk and Uncertainty* 14, 155-168.
- Camerer, C.F. and R.M. Hogarth, 1999. The effects of financial incentives in experiments: A review and capital-labor-production framework, *Journal of Risk and Uncertainty* 19, 7-42.
- Dohmen, T., A. Falk, D. Huffman, U. Sunde, J. Schupp and G.G. Wagner, 2005. *Individual risk attitudes: New evidence from a large, representative, experimentally-validated survey*, IZA Discussion Paper No. 1730.
- Edwards, W., 1953. Probability preferences in gambling, *The American Journal of Psychology* 66, 349-364.
- Etchart-Vincent, N. and O. l'Haridon, 2008. *Monetary Incentives in the Loss Domain and Behaviour toward Risk: An Experimental Comparison of Three Rewarding Schemes Including Real Losses*, Working Paper, HEC Business School.
- Faff, R., Mulino, D., and D. Chai, 2008. On the linkage between financial risk tolerance and risk aversion, *Journal of Financial Research* 31, 1-23.
- García-Gallego, A., N. Georgantzís, A. Jaramillo-Gutiérrez and M. Parravano, 2011. The SGG risk elicitation task: Implementation and results, The Papers 10/07, Department of Economic Theory and Economic History of the University of Granada.
- Harrison, G.W., Johnson, E., McInnes, M.M., Melayne, M. and E.E. Rutström, 2005. Temporal stability of estimates of risk aversion. *Applied Financial Economics Letters* 1, 31-35.
- Harrison, G.W., and E.E. Rutström, 2008. Risk aversion in the Laboratory. *Research in Experimental Economics*, 12, Bingley, UK: Emerald, 41-196.
- Harrison, G.W., Lau M.I. and E.E. Rutström, 2009. Risk attitudes, randomization to treatment, and self-selection into experiments, *Journal of Economic Behavior and Organization* 70, 498-507.
- Harrison, G. W. and E. Rutström, 2009. Expected utility theory and prospect theory: One wedding and a decent funeral. *Experimental Economics* 12, 133-158.
- Harrison, G. W., Moritz, S. and R. Pibernik, 2010. Characterizing risk attitudes of industrial managers. Working Paper 2010-01 Center for Economic Analysis of Risk.

Holt, C. A., and S. K. Laury, 2002. Risk aversion and incentive effects, *American Economic Review* 92, 1644-1655.

Holt, C. A., and S. K. Laury, 2005. Risk aversion and incentive effects: New data without order effects, *American Economic Review* 95, 902-912.

Kuhberger, A., M. Schulte-Mecklenbeck and J. Perner, 2002. Framing decisions: hypothetical and real', *Organizational Behavior and Human Decision Processes* 89, 1162-75.

Laury, S.K. and C.A. Holt, 2008. Further reflections on the reflection effect, *Research in Experimental Economics*, 12, Bingley, UK: Emerald, 405-440.

Sabater-Grande, G. and N. Georgantzís, 2002. Accounting for risk aversion in repeated prisoners' dilemma games: an experimental test. *Journal of Economic Behavior and Organization* 48, 37-50.

Smith, V. L., 1982. Microeconomic systems as an experimental science. *American Economic Review* 72, 923-955.

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

Wärneryd, K.E., 1996. Risk attitudes and risky behavior, *Journal of Economic Psychology* 17, 749-770.

Wiseman, D.B. and I.P. Levin, 1996. Comparing risky decision making under conditions of real and hypothetical consequences, *Organizational Behavior and Human Decision Processes* 66, 241-250.

Zuckerman, M., 1978. Sensation seeking, in: London H., Exner J., eds., *Dimension of personality*, John Wiley and Sons, New York.