Allais for all: Revisiting the paradox in a large representative sample

Steffen Huck · Wieland Müller

Abstract We administer the Allais paradox questions to both a representative sample of the Dutch population and to student subjects. Three treatments are implemented: one with the original high hypothetical payoffs, one with low hypothetical payoffs and a third with low real payoffs. Our key findings are: (i) violations in the non-lab sample are systematic and a large bulk of violations is likely to stem from non-familiarity with large payoffs, (ii) we can identify groups of the general population that have much higher than average violation rates; this concerns mainly the lowly educated and unemployed, and (iii) the relative treatment differences in the population at large are accurately predicted by the lab sample, but violation rates in all lab treatments are about 15 percentage points lower than in the corresponding non-lab treatments.

S. Huck
Department of Economics, University College London, Gower Street,
London WC1E 6BT, UK
e-mail: s.huck@ucl.ac.uk

S. Huck
WZB, Berlin, Germany

W. Müller (✉)
Department of Economics and VCEE, University of Vienna,
Brünnerstrasse 72, 1210 Vienna, Austria
e-mail: wieland.mueller@univie.ac.at

W. Müller
CentER, TILEC and Tilburg University, Tilburg, Netherlands

This paper presents evidence on the consistency of risk preferences with
expected utility theory in a representative population sample. We find that
consistency increases with task familiarity and is linked to several personal
characteristics such as education, income and asset holdings. Moreover, we
investigate the external validity of a laboratory experiment with a student
population that implemented the same choice problems as our household
panel study. We find that, in line with studies on other biases, deviations from
rationality observed in the lab provide a lower bound for deviations in the
population at large.

Recently, several studies have made significant progress in understanding
risk preferences in populations, making use of innovative survey methods and
field experiments (Harrison and List 2004) including game shows with large
stakes (Post et al. 2008; Andersen et al. 2008). From the perspective of these
studies, the present paper takes one step back by focussing on consistency of
risk preferences with expected utility theory in a representative subject pool—
well over 1,400 members of the CentER Panel, a representative sample of the
Dutch population. We do this by falling back on the oldest consistency test
of all—the Allais paradox (Allais 1953). Our results help to understand the
reliability and robustness of investigations into the actual distribution of risk
preferences in populations.

Our research strategy is threefold. First, we implement three different
treatments in the main experiment with the panel. We analyze the original
Allais question with payoffs of millions of Euros that, just as when Allais
asked Savage, were purely hypothetical. In our second treatment we scaled the
payments down but kept them hypothetical. Our third treatment used the same
downscaled payoffs but paid them out for real. This enables us to examine to
what extent violations are driven by lack of monetary incentives, on the one
hand, and non-familiarity with large sums of money on the other.

Second, we are able to exploit the wide range of background information
that is available for our subjects in order to study the roots of violations.[1]
Which personal characteristics are correlated with violations? Are violations

---

[1]Several other studies have also used the CentER panel as a subject pool. Let us briefly mention
some of these studies. Hey (2002) and Carbone (2005) analyze more complicated and sequen-
tial individual decision making tasks and do not find any background variable systematically
influencing behavior. Bellemare and Kröger (2007) study a trust game and find "that heterogeneity
in behavior is characterized by several asymmetries—men, the young and elderly, and low
educated individuals invest relatively less, but reward significantly more investments." (p. 183) von
Gaudecker et al. (2011a) elicit risk preferences and report that older people, women, the relatively
uneducated, and those with lower income are more risk averse. For another study on individual
risk attitudes using a large and representative German sample, see Dohmen et al. (2011).

a matter of insufficient education or limited experience with financial decision making? Can we identify 'problem groups' that are, perhaps, more likely to suffer (in particular late in life) from erroneous financial decision making?

Third, we conduct a laboratory experiment with the usual laboratory subject population (students) employing the same design that we used in the panel experiment. Thus, we are able to examine the external validity of a laboratory experiment in a clear and detailed manner. In particular, we can compare whether and how a lab study can tell us something about the population at large.

Pursuing our threefold research strategy we are, thus, able to present very detailed and comprehensive evidence on the Allais paradox. Our results are useful for several practical issues: (1) Our results point to a number of conditions that make standard theoretical predictions more likely to hold, (2) Our results identify certain parts of the population that, due to inconsistencies, may have difficulties in making sound financial decisions, and (3) Our results contribute to a better understanding of what can be reliably learned from laboratory experiments.

Along the first dimension of our research strategy we find that violations in the original paradox are likely to be driven by very high payoffs with which, in real life, virtually nobody has any practical experience. Violations in the original Allais problem are twice as high as in both downscaled versions. This effect has been observed before with student samples (Conlisk 1989); we show that the pattern extends to the general population and across socioeconomic characteristics. Perhaps this result is not surprising as it simply stresses that economic theory can be expected to work much better in environments with which agents have experience and are, thus, well-adapted. On the other hand, we find no substantial difference between the two downscaled versions. Whether subjects are incentivized or not, violations are much lower in *both* cases.[2]

Along the second dimension, we are able to identify a whole array of personal characteristics that correlate with inconsistent decision making. Education, occupation, income and asset holdings do all correlate with inconsistent decision making and in each case the direction of effects is as one would guess. The better educated are more consistent and so are those in employment, those who earn more and those who hold financial assets.

Finally, our methodological contribution reveals that the laboratory results are rather useful in predicting behavior in a general population. First, the relative treatment differences are precisely the same for both populations, panel and lab. Second, as demonstrated in a number of other studies (see Gächter et al. (2008) for a survey) the violations of standard theory observed

---

[2]For early studies of the Allais paradox see, e.g., MacCrimmon (1968), Slovic and Tversky (1974), Allais and Hagen (1979) and Kahneman and Tversky (1979). For the effect of downscaled payoffs see Conlisk (1989), Starmer and Sugden (1991), Harrison (1994), Burke et al. (1996), Fan (2002), and van de Kuilen and Wakker (2006).

in the lab provide a *lower bound* for violations observed in the population at large.[3]

The remainder of the paper is organized as follows. In Section 1, we describe the main characteristics of the CentERpanel and introduce the experimental design. In Section 2 we present our results obtained with the panel. We first give a quick overview of the results and then present a more detailed analysis, based on regression results, that also accounts for the effect of sociodemographic characteristics. In Section 3 we introduce our lab results and compare them to those obtained in the panel. Section 4 concludes.

1 Design and data collection

We administer the original "Allais questions," which consist of two pairwise lottery choices. Consider the following two choice problems. First, a subject is asked to choose between lotteries A and A^* where

$$A = \text{Certainty of } \text{€}1 \text{ Million} \quad \text{and} \quad A^* = \begin{cases} 1/100 \text{ Chance of } \text{€}0 \\ 89/100 \text{ Chance of } \text{€}1 \text{ Million} \\ 10/100 \text{ Chance of } \text{€}5 \text{ Million} \end{cases}$$

Second, a subject is asked to choose between lotteries B and B^* where

$$B = \begin{cases} 89/100 \text{ Chance of } \text{€}0 \\ 11/100 \text{ Chance of } \text{€}1 \text{ Million} \end{cases} \quad \text{and}$$

$$B^* = \begin{cases} 90/100 \text{ Chance of } \text{€}0 \\ 10/100 \text{ Chance of } \text{€}5 \text{ Million} \end{cases}$$

Of the four possible answers AB, A^*B^*, AB^*, and A^*B only the first two are consistent with expected utility theory (henceforth, EUT) whereas the last two are not.[4] Many laboratory experiments have shown that violations of EUT are frequent and that a larger share of subjects violating EUT chooses AB^* instead of A^*B.[5]

---

[3]Almost all of the experiments on the Allais paradox conducted so far have used students as their subjects. There are two notable exceptions. List and Haigh (2005) test the Allais paradox both with students and professional traders from the Chicago Board of Trade. They report that both students and professional traders show Allais paradox behavior, but find that traders do so to a smaller extent. Fatas et al. (2007) use students and politicians and report similar results with students being more prone to Allais paradox behavior.

[4]To see this note that by adding $0.89u(\text{€}0) - 0.89u(\text{€}1M)$ to both sides of the inequality $u(A) = u(\text{€}1M) > 0.01u(\text{€}0) + 0.89u(\text{€}1M) + 0.1u(\text{€}5M) = u(A^*)$ implies $u(B) = 0.89u(\text{€}0) + 0.11u(\text{€}1M) > 0.9u(\text{€}0) + 0.1u(\text{€}5M) + 0.1u(\text{€}5M) = u(B^*)$.

[5]See, e.g., MacCrimmon (1968), Slovic and Tversky (1974), Allais and Hagen (1979), Kahneman and Tversky (1979), Conlisk (1989), Starmer and Sugden (1991), Harrison (1994), Burke et al. (1996) and Fan (2002).

We have six simple treatments using a between-subjects design. To introduce these treatments, consider the following lotteries over three outcomes of monetary payoffs with probabilities as above, i.e., $A = (0, 1, 0)$, $A^* = (.01, .89, .10)$, $B = (.89, .11, 0)$, $B^* = (.90, 0, .10)$. Our three treatments were then as follows:

- Treatment HighHyp: Original Allais questions with high hypothetical payoffs of €0, €1 million, and €5 million.
- Treatment LowHyp: Allais questions with low hypothetical payoffs of €0, €5, and €25.
- Treatment LowReal: Allais questions with low real payoffs of €0, €5, and €25.

Note that the amounts of money we use in these treatments are the same as in Conlisk (1989) with the sole difference that he used dollars instead of euros. For all three treatments we had two sub treatments reversing the order of decisions. As we do not find any order effects in the data we pool the data throughout.

We collected data from a representative sample of the Dutch population. The experiments were conducted by CentERdata—an institute for applied economic and survey research for the social sciences—that is affiliated with Tilburg University in the Netherlands. CentERdata carries out its survey research mainly by using its own panel called CentERpanel. This panel is Internet based and consists of some 2000 households in the Netherlands which form a representative sample of the Dutch population.[6] One of the advantages of the CentERpanel is that the researcher has access to background information for each panel member such as demographic and financial data. Every weekend, the panel members complete a questionnaire on the Internet from their home.

After logging on to our experiment, panel members were randomly assigned to one of the six different treatments introduced above. After being informed about the nature of the experiment, subjects decided whether or not to participate—as common with many modules of the panel. For participating subjects, the next screen introduced an example of a pair of lotteries (which were referred to as "Options"). Subjects were told that their task would be to express preference for one of the two lotteries and, additionally, how the preferred lottery would be executed.[7] When subjects indicated that they were ready to start the experiment, they were, in two consecutive screens, presented with their two Allais questions. Only after answering both Allais questions, the two preferred lotteries were played out (by the computer) and subjects were informed about the outcome of their two preferred lotteries. In the treatments

---

[6]For more information about the CentERpanel and the way it is administered see http://www. uvt.nl/centerdata/en/whatwedo/thecenterpanel/.

[7]For more details see Appendix A which contains a translation of the screens used in the treatments with low payoffs. Note that the experiment was administered in Dutch.

with real monetary payments, subjects were paid according to the outcomes in both of their preferred lotteries.[8]

In total 1676 members of the CentERpanel logged on to our experiment. Of the subjects logging on, 1426 (85.1%) subjects decided to participate in our experiment while 250 (14.9%) subjects decided not to participate. Table 1 shows descriptive statistics of our sample. The column labeled "Participation" in Table 1 shows descriptive statistics of participating subjects in each of the three main treatments as well as statistics of subjects who chose not to participate in the experiment. The data in Table 1 is grouped according to gender, age, education, occupation and income. (The column labeled "Violation" shows statistics for participating subjects violating or not violating EUT, respectively, which we will analyze further below. It also contains tests on the role of socioeconomic characteristics for EUT violation which will also be discussed later.)

Concentrating on descriptive statistics for participating subjects in Table 1, we note that by and large most variables are relatively identically distributed across treatments. However, in some of the age and income brackets as well as in the category savings account, there is some more variation. A comparison of the descriptive statistics in the columns describing participating subjects with those of non-participating shows that there are no big differences except for the age categories. Basically, older people appear to be a little more reluctant to participate.

Since this causes concern about sample selection problems, we ran for all regressions reported below Heckman (1976) selection models using the variable "Ratio" as one of the exclusion variables. The variable "Ratio" measures the proportion of questionnaires completed by panel members in the three months proceeding our experiment. This variable can be assumed to affect the participation decision but not the decisions taken in the experiment. For none of the regressions we found evidence for a selection bias.[9]

2 Results

2.1 Descriptives

A summary of the experimental results is given in Table 2. The table shows both the absolute frequency of choices (left part) and the relative frequency of choices (right part). As mentioned in the introduction, we will concentrate

---

[8]Note the following about payments in treatment LowReal. CentERdata reimburses the telephone costs for filling in questionnaires by exchanging "CentERpoints" (1 CentERpoint = 0.01 Euro) to panel members' private bank accounts four times a year. Although lotteries were described in Euro amounts, subjects in the treatments with real monetary earnings were informed that: "In this experiment you can earn real money that will be paid in the form of CentERpoints."

[9]See Eckel and Grossman (2000), Bellemare and Kröger (2007), von Gaudecker et al. (2011b) and Harrison et al. (2009) for more evidence on selection issues.

Table 1 Descriptive statistics of the samples

Category		Participation				Violation		p-value, χ^2	
		YES			NO	NO	YES		
		HighHyp	LowHyp	LowReal					
Gender	Female	48.9	46.5	47.0	49.2	46.2	50.0	0.180	0.156
Age	Age 16–24	6.2	8.0	6.9	2.8	6.5	8.4		
	Age 25–34	24.4	19.2	19.7	8.0	20.6	21.5		
	Age 35–44	18.7	19.4	16.6	15.6	19.8	14.4		
	Age 45–54	17.5	19.2	23.3	24.8	20.2	20.1		
	Age 55–64	18.2	16.2	17.2	20.4	16.3	19.0		
	Age 65+	15.0	18.2	16.4	28.4	16.6	16.7		
Education	Primary education	5.0	6.0	5.7	7.6	5.0	7.1	0.008	
	Lower secondary education	28.5	26.6	25.2	32.0	25.0	30.2		
	Higher secondary education	12.0	13.4	13.7	11.6	13.9	11.4		
	Intermediate vocational training	20.0	20.2	20.2	18.8	19.7	21.3		
	Higher vocational training	24.2	23.2	23.5	21.2	23.9	22.9		
	University degree	10.2	10.6	11.6	8.8	12.6	7.1		
Occupation	Employed (contract)	54.4	49.1	52.3	41.6	54.7	45.2	0.003	
	Freelance or self-employed	3.7	3.2	3.8	4.8	4.0	2.5		
	Unemployed	1.5	2.4	1.9	1.2	1.5	3.0		
	Student	5.5	6.6	6.7	2.0	5.9	7.3		
	Works in own household	10.7	13.0	12.0	15.6	11.4	13.2		
	Retired	17.2	17.8	16.8	28.8	16.5	19.0		
	Other occupation	7.0	8.0	6.5	6.0	6.0	9.8		

Table 1 (continued)

Category		Participation				Violation		
		YES			NO	NO	YES	p-value, χ^2
		HighHyp	LowHyp	LowReal				
Household	HH gross income ≤ €2250	22.7	26.0	27.9	28.8	22.7	32.6	<0.001
Income	HH gross income €2251–€3130	28.4	21.8	23.5	22.0	24.0	24.9	
	HH gross income €3131–€4350	25.2	24.4	25.8	24.0	26.8	21.2	
	HH gross income ≥ €4351	23.7	28.0	22.9	25.2	26.5	21.2	
Assets	Holds assets	18.4	16.2	15.5	12.0	18.0	13.2	0.025
Savings acct.	Has savings account	59.4	53.5	50.6	50.0	52.4	57.8	0.062
	Maximum no. of observations	401	501	524	250	988	438	

Notes: HighHyp stands for high hypothetical payoffs, LowHyp for low hypothetical payoffs, and LowReal for low real payoffs. Except for the number of observations, numbers indicate column percentages. Since some of the members of the CentERpanel did not complete the Dutch Household Survey, some observations are missing. The column labeled "p-value, χ^2" shows p-levels of χ^2 tests for differences between proportions of violating and non-violating subjects in the category listed in column 1

Table 2 Summary of experimental results in the panel

Treatment	Absolute frequency of choices					Relative frequency of choices				
	AB	A^*B^*	AB^*	A^*B	Σ	AB	A^*B^*	AB^*	A^*B	Violations
HighHyp	82	121	136	62	401	20.4	30.2	33.9	15.5	49.4
LowHyp	22	373	77	29	501	4.4	74.4	15.4	5.8	21.2
LowReal	22	368	97	37	524	4.2	70.2	18.5	7.1	25.6
Σ	126	862	310	128	1426	8.8	60.4	21.7	9.0	30.7

Note: HighHyp stands for high hypothetical payoffs, LowHyp for low hypothetical payoffs, and LowReal for low real payoffs

our analysis on the incidence of subjects' EUT violation in all treatments. However, we will also shortly answer the question whether violations, once they occur, are systematic.

Violation of EUT Note that the right-most column in Table 2 indicates that violations of EUT are observed in all treatments. In fact, we observe 49.5%, 19.6% and 25.6% violations of EUT in treatments HighHyp, LowHyp, and LowReal, respectively. Furthermore, in all treatments we observe that the fraction of EUT-violating AB^* answers is higher than the fraction of EUT-violating A^*B answers. The Z-statistic proposed in Conlisk (1989) indicates that the first fraction is significantly higher than the latter fraction at $p < 0.001$ in all treatments. An interesting question we can answer with our data is whether the differences we report here for the aggregate data are "general" in the sense of applying across socioeconomic attributes or whether they are driven by only some of those attributes. The answer is provided in Tables 5, 6, 7 and 8 in Appendix B, which are structured as Table 1 and provide—for all data and for the three treatments separately—the relative frequency of choices for subjects with various socioeconomic attributes. We observe that EUT violations occur across all socioeconomic attributes and that the "Allais" pattern of more AB^* violations than A^*B violations is significant for most socioeconomic attributes in all treatments (see the column labeled "Sign. of Conlisk's Z-statistic" in Tables 5–8 in Appendix B). We conclude that, as in earlier studies, violations of EUT are observed and that they are systematic in the sense that AB^* is chosen more often than A^*B, mostly independent of socioeconomic background characteristics. To facilitate comparison, note that Conlisk (1989) using a student sample for his "Basic Version" (which is comparable to our treatment HighHyp) reports the following relative frequencies of AB, A^*B^*, AB^*, and A^*B choices: 7.6%, 41.9%, 43.6%, and 6.8%. Thus, he observes EUT violation in 50.4% of the cases which compares to 45.5% in our panel treatment HighHyp.

The effect of high versus small hypothetical payoffs Next consider the effect of high versus small hypothetical payoffs on the extent of EUT violation. For this purpose we compare the rates of EUT violations in treatments HighHyp and LowHyp. Table 2 shows that the rate of EUT violations drops from

49.4% in treatment HighHyp to 19.6% in treatment LowHyp. The D-statistic proposed in Conlisk (1989) indicates that this difference is highly significant at $p < 0.0001$ ($D = 9.115$). Inspecting the relative frequencies of choices in Table 2 shows that moving from HighHyp to LowHyp sharply increases the fraction of choices consistent with expected value maximization (A^*B^*) at the expense of all other three possible responses. In particular, many more subjects prefer the payoff-maximizing choice A^* over A when (hypothetical) payoffs become small. A possible explanation of this result is due to the fact that subjects in treatment LowHyp can be expected to be more familiar with the lower amounts of money leading them to make fewer mistakes.[10] Again, with our data we can check whether the result regarding the effect of varying the (hypothetical) stake size just shown for the aggregate data also applies when the data is broken down to various socioeconomic characteristics. Column 3 labeled "Significance of Conlisk's D-statistic HighHyp vs LowHyp" in Table 9 in Appendix B shows that the answer to this question is, with a few exceptions, yes.

The effect of (small) real versus (small) hypothetical payoffs   Finally, consider the effect of (small) real versus (small) hypothetical payoffs on the extent of EUT violation. To analyze this, compare the rates of EUT violation in treatments LowHyp and LowReal. Table 2 shows that the rate of EUT violations is 19.6% in LowHyp whereas it is 25.6% in treatment LowReal. Thus, we see a slight increase in the share of EUT violations when we move from (small) hypothetical to (small) real payoffs. The D-statistic in Conlisk (1989) indicates that this difference is significant ($D = -1.6716$, $p = 0.047$). In contrast, Harrison (1994) and Burke et al. (1996) report that the use of low real instead of low hypothetical payoffs reduces the extent of EUT violation. For a broader overview on how incentives affect behavior in decisions under risk, see Camerer (1995, p. 634f). Note that the result regarding the switch from (small) hypothetical to (small) real payoffs on the extent of EUT violation is usually not significant when one zooms in on socioeconomic characteristics, as shown in column 4 labeled "Significance of Conlisk's D-statistic LowHyp vs LowReal" in Table 9 in Appendix B.

Note that our results concerning the extent of EUT violation and the effect of high versus small hypothetical payoffs are not entirely new. We show, however, that they extend to a general population and across socioeconomic characteristics. This should be of interest due to the current discussion about the relationship between results obtained in the lab and those obtained in other settings (see, e.g., Levitt and List 2007).

---

[10]Conlisk (1989) points out that this effect is in line with (a) Machina's (1982) fanning out model that predicts Allais behavior for large payoffs and (b) the observation that EUT converges to expected payoff maximization for small payoffs. Notice, however, that this consistency argument is not an explanation—for it leaves open why fanning occurs and is more dramatic in its consequences with high payoffs. Non-familiarity with high payoffs is such an explanation and may, in fact, be adequately captured by fanning out of indifference curves.

Let us now turn to providing answers to the first of the two new and main dimensions of our research strategy by inspecting the role of socioeconomic background variables in subjects' behavioral responses to the Allais questions. Refer to Table 1 that under the heading "Violation" shows descriptive statistics of the subsamples violating and not violating EUT as well as p-levels of χ^2 tests. (For the latter, see the notes below Table 1.) Regarding gender, Table 1 reveals that women are slightly more likely to violate EUT than men. With respect to age, Table 1 does not suggest a clear effect although we note that the age bracket's [35–44] relative share is higher in the panel's subpopulation not violating EUT. Regarding education levels, those with lower secondary education and those subjects with a university degree stand out somewhat in the panel. The former because they violate EUT more often and the latter because they violate EUT less often. The most noticeable effect regarding occupation is that those employed on a contractual basis have a higher relative share in the subsample not violating EUT. Finally, with respect to household income, Table 1 does not suggest a clear effect.

Moreover, refer to the rightmost column labeled "p-value, χ^2" in Table 1 that shows p-levels of χ^2 tests for differences between proportions of violating and non-violating subjects in the category listed in column 1.[11] The χ^2 tests indicate the strongest differences in violation behavior in the categories of education, occupation and household income.

2.2 Econometrics and the role of socioeconomic characteristics

To test for across-treatment differences controlling for subjects' sociodemographic characteristics and to check whether any of these characteristics are correlated with behavior, we ran probit regressions with the variable "Violate" as the dependent variable. "Violate" is equal to 1 if a subject's answer to the Allais questions violates EUT (i.e., answers A^*B or AB^*), and is equal to 0 otherwise (i.e., answers AB or A^*B^*). The background variables we include in the regression are the ones shown in Table 1 above. The results are shown in Table 3 which reports marginal effects. Regression (1) includes all data whereas regressions (2) to (4) show results for each of the three treatments separately. Recall from the end of Section 2 that we did not find evidence for a selection bias due to non-response.

Let us first briefly reconsider across-treatment differences. For this purpose, refer to regression (1) in Table 3 which includes all data and controls for background variables. Importantly, note that in regression (1) the omitted treatment dummy is the one for LowHyp. Inspecting the treatment coefficients, we note that the coefficient for HighHyp is positive and big (0.302) and highly statistically significant whereas the coefficient of LowReal is also positive (0.053) but rather small and only borderline significant.

---

[11]Note that for the multinomial categories in the leftmost column in Table 1, the χ^2-tests check for the joint hypothesis that the violation rates are identical across all categories.

Table 3 Results of probit regressions on violation of EUT

	(1) All data	(2) Only HighHyp	(3) Only LowHyp	(4) Only LowReal
HighHyp	0.307*** (9.22)			
LowReal	0.058* (1.87)			
Female	0.007 (0.25)	−0.074 (1.25)	0.056 (1.33)	−0.002 (0.05)
Age 25–34	−0.062 (0.77)	0.005 (0.03)	0.017 (0.15)	−0.200* (1.77)
Age 35–44	−0.115 (1.45)	−0.186 (1.07)	−0.027 (0.24)	−0.175 (1.46)
Age 45–54	−0.050 (0.61)	−0.082 (0.46)	−0.011 (0.09)	−0.125 (0.96)
Age 55–64	−0.072 (0.87)	−0.162 (0.89)	0.031 (0.25)	−0.152 (1.21)
Age 65+	−0.145 (1.64)	−0.157 (0.75)	−0.134 (1.21)	−0.125 (0.86)
Lower second. edu.	−0.035 (0.63)	0.177 (1.35)	−0.147** (2.23)	0.045 (0.51)
Higher second. edu.	−0.086 (1.50)	0.150 (1.07)	−0.152** (2.40)	−0.027 (0.30)
Intermed. voc. training	−0.025 (0.43)	0.172 (1.26)	−0.134** (1.97)	0.116 (1.19)
Higher voc. training	−0.034 (0.59)	0.288** (2.18)	−0.139** (1.98)	−0.010 (0.11)
University degree	−0.134** (2.21)	0.186 (1.28)	−0.188*** (2.85)	−0.152* (1.69)
Employed (contract)	−0.130** (2.53)	−0.234** (2.04)	−0.127* (1.82)	−0.118 (1.50)
Freelance or self-empl.	−0.149** (2.11)	−0.435*** (3.08)	−0.081 (0.74)	−0.025 (0.21)
Unemployed	0.065 (0.65)	−0.023 (0.09)	0.043 (0.34)	0.046 (0.30)
Student	−0.101 (1.16)	−0.156 (0.79)	−0.033 (0.26)	−0.169 (1.43)
Works in own household	−0.061 (1.10)	−0.224 (1.78)	−0.015 (0.19)	−0.068 (0.82)
Retired	0.013 (0.19)	−0.063 (0.43)	0.145 (1.45)	−0.145 (1.56)
HH gr. inc. €2251–€3130	−0.077** (2.26)	−0.042 (0.54)	−0.047 (0.95)	−0.129*** (2.64)
HH gr. inc. €3131–€4350	−0.110*** (3.23)	−0.167** (2.10)	−0.033 (0.67)	−0.138*** (2.84)
HH gr. inc. ≥ €4351	−0.083** (2.28)	−0.073 (0.88)	−0.076 (1.49)	−0.089* (1.69)
Assets	−0.081** (2.29)	−0.118 (1.59)	0.001 (0.02)	−0.112** (2.02)
Savings account	0.054** (2.05)	0.093 (1.64)	0.043 (1.13)	0.021 (0.53)
No. of observations	1424	400	500	524

Marginal effects are reported

Notes: HighHyp stands for high hypothetical payoffs, LowHyp for low hypothetical payoffs, and LowReal for low real payoffs. Absolute value of z-statistics in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Omitted categories are treatment LowHyp; age interval [16–24]; primary education; "other" occupation; household gross income smaller or equal to €2250. Two occupation observations missing

To analyze the effect of socioeconomic background variables economet-
rically, we examine regression (1) in Table 3. We make the following
observations.

- Controlling for other characteristics, gender and age have no significant
  influence on the extent of EUT violation.[12]
- Regarding education, we find a strong tendency for violations to be
  reduced with further education.[13] Overall, there is a strong effect of higher
  education that also shows in the separate specifications for both treatments
  with low payoffs. In LowHyp everything that improves on primary educa-
  tion goes hand in hand with reduced violations. Only in HighHyp there
  is no effect of education. This suggests an interesting interaction effect of
  experience with a decision domain and education. In the absence of any
  experience (as in HighHyp) education on its own does little to improve
  performance. Only if coupled with experience education is aligned with
  consistency.
- Of the various occupational affiliations listed in Table 3, we find that the
  unemployed and 'others' do much worse than the employed, self-employed
  and freelancers.[14] This is more pronounced in treatments with hypothetical
  payoffs.
- Regarding income, we notice that having a higher gross monthly household
  income (vis-à-vis the control group with the lowest gross monthly house-
  hold income) goes along with reduced EUT violations.[15] Interestingly,
  this is particularly pronounced in the treatment LowReal when actual
  money is at stake. (One could have conjectured that it would be the other
  way round as the marginal utility of making some money and, hence, the
  incentive to think a little harder might be higher for those on low incomes.
  Alas, it does not work this way.)
- Finally, subjects holding assets have significantly lower EUT violations
  (by about 8%) whereas subjects with a savings account have significantly
  higher EUT violations (by about 5%). Maybe not surprisingly, subjects
  holding assets tend to be expected value maximizers (mainly choosing
  A^*B^*) while subjects who only have a savings account display "Allais"
  behavior tending toward the choice of AB^*.[16]

---

[12]In light of recent findings about sharply declining numeracy skills in the (British) population
above 55 (Banks 2006) this is perhaps slightly surprising.

[13]Wald tests indicate, however, that the effects of the education levels below university degree
listed in Table 3 are not statistically different.

[14]A Wald test indicates that the effect of these two occupations is not statistically different.

[15]Wald tests indicate that the effects of the three income variables listed in Table 3 are not
statistically different. Furthermore, controlling for household size leaves the regression results
reported in Table 3 virtually unchanged.

[16]To look at the effect of holding assets or a savings account more closely, we defined the variable
"only assets" which equals 1 if a subject holds assets but has no savings account (otherwise it
equals 0), the variable "only savings account" which equals 1 if a subject has a savings account
but holds no assets (otherwise it equals 0), and the variable "assets & savings account" which

In all a picture emerges that is reminiscent of recent studies by Benjamin et al. (2006), Burks et al. (2009) and Dohmen et al. (2010) who show that a range of behavioral biases are correlated with (or may even stem from) cognitive limitations and low IQ. We find that violations are more prevalent in those who are lowly educated, unemployed, on low income, and who have no significant asset holdings. This is, of course, particularly worrying as imprudent financial decision making and bad planning for retirement has the worst consequences in that group.

In Appendix C we complement the above analysis by running multinomial logit regressions using all four answers AB, A^*B^*, AB^*, and A^*B, and choosing the answer representing expected value maximization, A^*B^*, as the base outcome. The results (whose interpretation is less straightforward) are shown in Tables 10, 11, 12 and 13.

3 The lab experiment

As mentioned in the introduction, the third dimension of our research strategy is concerned with the external validity of laboratory experiments that are typically carried out with rather homogenous subject pools. Of course, the preceding section has shown that there are important sources of heterogeneity in the population at large that simply cannot be detected when the subject pool is restricted to students. The same is, of course, true for any highly selected convenience sample. But what about the questions we analyzed first—the effects of different treatments, the differences between high and low and real and hypothetical payoffs? Would a lab experiment give us reliable results to analyze such questions (as it has been implicitly assumed for a long time in the experimental community, perhaps negligently without much testing)? To shed more light on these issues we conducted an additional lab experiment in the laboratory of Tilburg University using Dutch speaking student subjects drawn from the normal subject pool.

The lab experiment was conducted in the same way as the experiment using the CentERpanel. That is, student subjects did the experiment using a web browser in the lab and using the same screens as the subjects in the panel. However, there were two small exceptions. First, lab subjects received a 10 Euro show-up fee. (Potential participants were informed about this in

---

equals 1 if a subject holds assets and has a savings account (otherwise it equals 0). Hence, the reference group consists of those subjects who neither hold assets nor have a savings account. Replacing the variables "assets" and "savings account" in regression (1) in Table 3 by the new variables "only assets," "only savings account," and "assets & savings account," leaves the other variables of regression (1) almost unchanged (including significance levels) and shows that while the coefficients of the variables "only assets" and "assets & savings account" are negative (−0.073 and −0.032) but insignificant, the coefficient of the variable "only savings account" is positive (0.055) and significant at the 5% level. So it is not only the financially savvy who hold assets who do comparatively well but also people without any savings—perhaps because, having no financial cushion, they cannot afford making many mistakes.

Table 4 Summary of experimental results in the lab

Treatment	Absolute frequency of choices					Relative frequency of choices				
	AB	A^*B^*	AB^*	A^*B	Σ	AB	A^*B^*	AB^*	A^*B	Violations
HIGHHYP	4	41	20	5	70	5.7	58.6	28.6	7.1	35.7
LOWHYP	0	75	4	0	79	0	94.9	5.1	0	5.1
LOWREAL	1	67	5	1	74	1.4	90.5	6.8	1.4	8.2
Σ	5	183	29	6	223	2.2	82.1	13.0	2.7	15.7

Notes: HighHyp stands for high hypothetical payoffs, LowHyp for low hypothetical payoffs, and LowReal for low real payoffs

the invitation E-mail.) But of course, mirroring the panel design again, only subjects assigned to treatments with real payment had the chance to earn additional money during the experiment. This was not announced prior to the experiment. Second, lab subjects were not offered the choice of not participating in the experiment once they had reported to the lab and the experiment was started. This was done in an effort to mimic the normal procedures in lab experiments where by reporting to the lab, a subject usually confirms his or her decision to participate. Note that when we move from the panel to the lab sample, both the subject pool and the environment changes. We deliberately accepted these two simultaneous changes as our aim was to contrast the results obtained in the panel with those obtained in a normal lab experiment.[17]
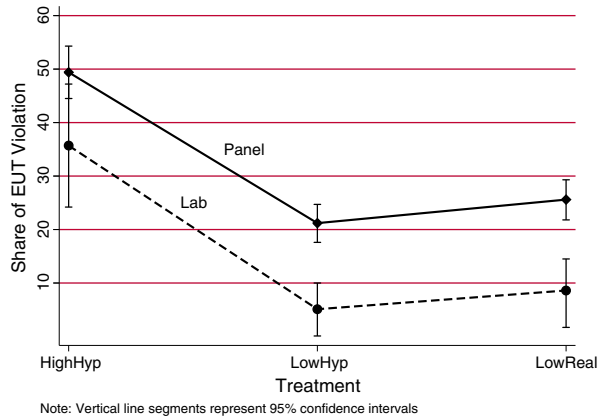
After the experiment we asked subjects to fill in a questionnaire in which we elicited some basic background information. Naturally, the information we collected from lab subjects is very limited and cannot be compared in scope and quality to the background information available from members of CentERpanel. The lab experiments were conducted in December 2006 using 223 subjects in total.

As in the panel experiment we did not observe any order effects of presenting the Allais questions, so we present only pooled data in Table 4 which shows the same information for the lab data that Table 2 showed for the panel. We make the following observations. First, as in the panel experiments, we observe EUT violations in all treatments, although to a much lesser degree.[18] This mirrors the main result in Gächter et al.'s (2008) meta-study: Violations from orthodox theoretical predictions and biases observed in the lab form a lower bound for violations and biases observed in the population at large. Second, as in the panel, moving from high hypothetical payoffs to low hypothetical payoffs reduces the extent of EUT violation significantly

---

[17] von Gaudecker et al. (2011b) offer an analysis of the individual effects of implementation mode and of subject pool selection in a risk preference elicitation study and find that differences in behavior are due to selection and not implementation mode.

[18] Again, we observe that the fraction of EUT-violating AB^* answers is significantly higher than the fraction of EUT-violating A^*B answers in all lab treatments ($p < 0.001$, Conlisk's (1989) Z-statistic).

Note: Vertical line segments represent 95% confidence intervals

($p < 0.001,\ D = 4.881$). Third, moving from low hypothetical payoffs to low real payoffs increases the extent of EUT violation slightly but insignificantly ($p < 0.226,\ D = -0.7525$). The similarities between the observations in the panel and in the lab are evident.

Figure 1 shows the shares of choices violating EUT in the two subsamples. It appears that the graph indicating the share of EUT violation in the panel can quite accurately be obtained by shifting the graph indicating the share of EUT violation in the lab upwards by about 15 percentage points.[19] This means that although the share of EUT violations is consistently higher in the panel than in the lab, the comparative statics results of moving from one treatment to another could have been reliably predicted by the lab experiments.

4 Conclusions

Using a representative sample of the Dutch population we revisit the Allais paradox. Our main results are threefold. First, as in previous lab samples, the violations of EUT are systematic in the population at large and much lower when stakes are low. Second, there is considerable heterogeneity in the population and violations are particularly prevalent among the lowly educated, those poor in income and asset holdings, and the unemployed. Third, comparing the panel results with a laboratory experiment we find that the relative treatment differences are identical in the panel and the lab but violation rates in all lab treatments are about 15 percentage points lower than in the corresponding non-lab treatment.

Our findings appear to imply two general messages. First, laboratory experiments with convenience samples of students might be more useful to study

---

[19]The difference in the extent of EUT violation between the panel and the lab is significant for all three treatments (HighHyp: $p = 0.014,\ D = -2.1732$; LowHyp: $p < 0.001$; $D = -5.2220$; LowReal: $p < 0.001,\ D = -4.6935$).

relative effects rather than *absolute* levels (see also Levitt and List (2007) who make a similar point in the context of social preferences). When it comes to the absolute measurement of behavior, it appears that lab results will draw a too optimistic picture. The population at large, it turns out, is less consistent with EUT than student samples are. Second, our results suggest that the predictive power of EUT in a general population is correlated with socioeconomic characteristics. In particular, parts of the population that are more likely to experience economic hardship are less consistent.

Of course, there exists a large literature on non-expected utility theories such as Kahneman and Tversky's (1979) prospect theory or Machina's (1982) fanning-out theory (both of which can explain the Allais paradox) or Viscusi's (1989) prospective reference theory which predicts the paradox. Earlier laboratory experiments (see Camerer (1995) or Starmer (2000) for surveys) have documented the Allais paradox in student samples. Our paper highlights that, if anything, these studies underestimate the true prevalence of the paradox in general populations and indicates how violations are correlated with observable characteristics.

Appendix A: Instructions (Translation)

The experiment was administered in Dutch. Here we give a translation of the screens presented in treatment LowHyp and [LowReal]

Screen 1:
This research is conducted by researchers of Tilburg University and University College London. The questionnaire consists of two choice problems in which you will be asked to make a choice between two situations. Based on your choices and luck you may win an amount of money. Please note: In this experiment all amounts are hypothetical, in reality you cannot win any money. [In LowReal: In this experiment you can earn real money that will be paid in the form of CentERpoints.]

If you do not want to participate as a matter of principle, you can indicate this below. You will then go directly to the end of the questionnaire.

○ I continue with the questionnaire.
○ No, I do not want to participate in this questionnaire.

Continue

Screen 2:
Instructions

You will shortly be presented with two questions. You will be asked to make a choice between two options which provide you with different chances to win something. Please see an example of such a situation here below. In the first option you have a chance of 80% to win nothing and a chance of 20% to win 10 Euro. The second option provides you with a chance of 20% of nothing and a chance of 80% of winning 20 Euro.

OPTION 1:	80% chance	nothing	(if number is between 1 and 80) and
	20% chance	10 euro	(if number is between 81 and 100)
OPTION 2:	20% chance	nothing	(if number is between 1 and 20) and
	80% chance	20 euro	(if number is between 21 and 100)

We would like to know whether you prefer Option 1 or Option 2 (in these instructions you don't have to choose yet). After you have made the choice, the computer will play out the option you chose. The computer generates a random number that is between 1 and 100. The chance distribution of the chosen option then defines how much you win with this number.

For example: in the Option 1 above you get nothing if the computer generates a number between 1 and 80 (this is indicated above in Option 1 in brackets), but if the computer generates a number between 81 and 100 you will get 10 Euro. In Option 2 you get nothing if the computer generates a number between 1 and 20, but with a number between 21 and 100 you win 20 Euro. As already mentioned, it concerns hypothetical amounts here, in reality you cannot win any money. [In LowReal: If you win something then this amount will be added to your account of CentERpoints.]

If you are ready to start the experiment, press "Continue."

Continue

Screen 3:
Which of the following two options do you prefer?

OPTION A:	certainty	of 5 euro	(if number is between 1 and 100) and
OPTION B:	1% chance	of nothing	(if number is 1) and
	89% chance	of 5 euro	(if number is between 2 and 90) and
	10% chance	of 25 euro	(if number is between 91 and 100)

○ Option A
○ Option B
[Continue]

Screen 4:
Which of the following two options do you prefer?

OPTION C:	89% chance	of nothing	(if number is between 1 and 89) and
	11% chance	of 5 euro	(if number is between 90 and 100)
OPTION D:	90% chance	of nothing	(if number is between 1 and 90) and
	10% chance	of 25 euro	(if number is between 91 and 100)

○ Option C
○ Option D
[Continue]

Screen 5:
You have now made the two decisions. Press "Continue" to see the results of the options you chose.

[Continue]

Screen 6:
In the first question (option A or B) you have chosen Option X ([description of the chosen option]). The computer generated the number [random number]. Thus, you have won [in treatment LowHyp: the hypothetical] amount of [...] euro with this option.

In the second question (option C or D) you have chosen Option Y ([description of the chosen option]). The computer generated the number [random number]. Thus, you have won the [in treatment LowHyp: the hypothetical] amount of [...] euro with this option.

In total you have won the [in treatment LowHyp: the hypothetical] amount of [...] euro in this experiment.

[Continue]

Screen 7:
Do you have any comments regarding the questionnaire?

○ Yes
○ No
[Continue]

Screen 8 [In case the answer to the question on Screen 7 was Yes.] **:**
You can type in your comments below.

Table 5 Relative frequency of choices: all data

All treatments		Relative freq. of choices				# obs.	Sign. of Conlisk's Z-statistic	Sign. of χ² test
Category		AB	A*B*	AB*	A*B			
Gender	Female	11.7	55.9	23.7	8.7	675	***	***
	Male	6.3	64.6	20.0	9.2	751	***	
Age	Age 16–24	5.0	58.4	22.8	13.9	101	*	n.s.
	Age 25–34	8.8	59.6	21.5	10.1	297	***	
	Age 35–44	8.1	67.6	18.5	5.8	259	***	
	Age 45–54	8.3	61.1	21.9	8.7	288	***	
	Age 55–64	11.9	54.1	24.6	9.4	244	***	
	Age 65+	8.9	60.3	21.9	8.9	237	***	
Education	Primary education	12.5	48.8	27.5	11.3	80	***	***
	Lower secondary education	11.9	53.3	23.0	11.9	379	***	
	Higher secondary education	9.1	64.2	17.1	9.6	187	**	
	Intermediate vocational training	9.1	58.5	23.3	9.1	287	***	
	Higher vocational training	6.3	64.0	21.7	8.0	336	***	
	University degree	4.5	75.5	18.7	1.3	155	***	
Occupation	Employed (contract)	8.7	64.5	19.2	7.6	738	***	***
	Freelance or self-employed	13.7	64.7	17.6	3.9	51	**	
	Unemployed	0.0	53.6	28.6	17.9	28	n.s.	
	Student	3.3	61.1	24.4	11.1	90	**	
	Works in own household	14.0	52.0	24.0	9.9	171	***	
	Retired	6.5	59.8	23.2	10.6	246	***	
	Other occupation	11.8	46.1	30.4	11.8	102	***	

Table 5 (continued)

| All treatments | Relative freq. of choices | | | | # obs. | Sign. of Conlisk's | Sign. of |
Category	AB	$A*B*$	$AB*$	$A*B$		Z-statistic	χ^2 test
Household							
HH gross income $\leq$ €2250	9.5	51.5	25.1	13.9	367	***	***
Income							
HH gross income €2251–€3130	11.6	56.9	24.0	7.5	346	***	
HH gross income €3131–€4350	8.1	65.9	16.8	9.2	358	***	
HH gross income $\geq$ €4351	6.2	67.6	21.1	5.1	355	***	
Holds assets							
Yes	8.5	66.9	18.6	5.9	1,190	***	**
No	8.9	59.2	22.4	9.6	236	***	
Has savings							
Yes	8.0	59.1	23.5	9.3	655	***	n.s.
Account							
No	9.8	62.0	19.7	8.5	771	***	

Notes: The column labeled "Sign. of Conlisk's Z-statistic" indicates significance for the test that violations of EUT are systematic (more $AB*$ than $A*B$ violations). The column labeled "Sign. of χ^2 test" indicates significance levels of χ^2 tests for differences between proportions of choices in the category listed in column 1. ***, **, * indicates significance at the 1%, 5% and 10% level, respectively; *n.s.* indicates non-significance

Table 6 Relative frequency of choices: treatment HighHyp

Treatment HighHyp		Relative freq. of choices				# obs.	Sign. of Conlisk's Z-statistic	Sign. of χ^2 test
Category		AB	A*B*	AB*	A*B			
Gender	Female	26.5	26.0	33.2	14.3	196	***	**
	Male	14.6	34.1	34.6	16.6	205	***	
Age	Age 16–24	20.0	24.0	32.0	24.0	25	n.s.	n.s.
	Age 25–34	22.4	23.5	36.7	17.3	98	***	
	Age 35–44	16.0	45.3	30.7	8.0	75	***	
	Age 45–54	20.0	31.4	31.4	17.1	70	**	
	Age 55–64	27.4	26.0	34.2	12.3	73	***	
	Age 65+	15.0	28.3	36.7	20.0	60	**	
Education	Primary education	35.0	25.0	35.0	5.0	20	***	n.s.
	Lower secondary education	19.3	28.9	28.1	23.7	114	n.s.	
	Higher secondary education	22.9	33.3	29.2	14.6	48	*	
	Intermediate vocational training	23.8	31.3	33.8	11.3	80	***	
	Higher vocational training	16.5	27.8	39.2	16.5	97	***	
	University degree	17.1	34.1	43.9	4.9	41	***	
Occupation	Employed (contract)	20.2	32.6	33.0	14.2	218	***	n.s.
	Freelance or self-employed	40.0	40.0	13.3	6.7	15	n.s.	
	Unemployed	0.0	33.3	33.3	33.3	6	n.s.	
	Student	13.6	31.8	40.9	13.6	22	**	
	Works in own household	32.6	25.6	30.2	11.6	43	**	
	Retired	15.9	26.1	36.2	21.7	69	*	
	Other occupation	14.3	21.4	46.4	17.9	28	**	

Table 6 (continued)

Treatment HighHyp		Relative freq. of choices				# obs.	Sign. of Conlisk's	Sign. of
Category		AB	A^*B^*	AB^*	A^*B		Z-statistic	χ^2 test
Household	HH gross income ≤ €2250	18.7	25.3	35.2	20.9	91	*	*
Income	HH gross income €2251–€3130	25.4	21.9	38.6	14.0	114	***	
	HH gross income €3131–€4350	20.8	38.6	24.8	15.8	101	*	
	HH gross income ≥ €4351	15.8	35.8	36.8	11.6	95	***	
Holds assets	Yes	21.6	35.1	32.4	10.8	327	***	n.s.
	No	20.2	29.1	34.3	16.5	74	***	
Has savings	Yes	18.9	29.4	34.9	16.8	163	***	n.s.
Account	No	22.7	31.3	32.5	13.5	238	***	

Notes: HighHyp stands for high hypothetical payoffs. The column labeled "Sign. of Conlisk's Z-statistic" indicates significance for the test that violations of EUT are systematic (more AB^* than A^*B violations). The column labeled "Sign. of χ^2 test" indicates significance levels of χ^2 tests for differences between proportions of choices in the category listed in column 1. ***, **, * indicates significance at the 1%, 5% and 10% level, respectively; n.s. indicates non-significance

Table 7 Relative frequency of choices: treatment LowHyp

Treatment LowHyp Category		Relative freq. of choices				# obs.	Sign. of Conlisk's Z-statistic	Sign. of χ^2 test
		AB	A^*B^*	AB^*	A^*B			
Gender	Female	5.6	69.5	18.9	6.0	233	***	*
	Male	3.4	78.7	12.3	5.6	268	***	
Age	Age 16–24	0.0	72.5	17.5	10.0	40	n.s.	n.s.
	Age 25–34	3.1	77.1	11.5	8.3	96	n.s.	
	Age 35–44	5.2	79.4	14.4	1.0	97	***	
	Age 45–54	4.2	78.1	12.5	5.2	96	**	
	Age 55–64	7.4	61.7	23.5	7.4	81	***	
	Age 65+	4.4	74.7	15.4	5.5	91	**	
Education	Primary education	6.7	46.7	33.3	13.3	30	**	***
	Lower secondary education	8.3	67.7	17.3	6.8	133	***	
	Higher secondary education	6.0	76.1	10.4	7.5	67	n.s.	
	Intermediate vocational training	4.0	75.2	14.9	5.9	101	**	
	Higher vocational training	0.9	81.0	14.7	3.4	116	***	
	University degree	0.0	90.6	9.4	0.0	53	***	
Occupation	Employed (contract)	3.7	82.5	10.2	3.7	246	***	***
	Freelance or self-employed	6.3	81.3	12.5	0.0	16	*	
	Unemployed	0.0	66.7	25.0	8.3	12	n.s.	
	Student	0.0	69.7	18.2	12.1	33	n.s.	
	Works in own household	6.2	63.1	23.1	7.7	65	**	
	Retired	1.1	70.8	20.2	7.9	89	**	
	Other occupation	17.5	55.0	20.0	7.5	40	*	

Table 7 (continued)

Treatment LowHyp		Relative freq. of choices				# obs.	Sign. of Conlisk's	Sign. of
Category		AB	A^*B^*	AB^*	A^*B		Z-statistic	χ^2 test
Household	HH gross income $\leq$ €2250	7.7	63.8	18.5	10.0	130	**	***
Income	HH gross income €2251–€3130	6.4	73.4	16.5	3.7	109	***	
	HH gross income €3131–€4350	2.5	76.2	13.9	7.4	122	*	
	HH gross income $\geq$ €4351	1.4	83.6	12.9	2.1	140	***	
Holds assets	Yes	1.2	81.5	14.8	2.5	420	***	n.s.
	No	5.0	73.1	15.5	6.4	81	***	
Has savings	Yes	3.0	74.3	18.3	4.5	233	**	*
Account	No	6.0	74.7	12.0	7.3	268	***	

Notes: LowHyp stands for low hypothetical payoffs. The column labeled "Sign. of Conlisk's Z-statistic" indicates significance for the test that violations of EUT are systematic (more AB^* than A^*B violations). The column labeled "Sign. of χ^2 test" indicates significance levels of χ^2 tests for differences between proportions of choices in the category listed in column 1. ***, **, *, indicates significance at the 1%, 5% and 10% level, respectively; n.s. indicates non-significance

Table 8 Relative frequency of choices: treatment LowREAL

Treatment LowREAL		Relative freq. of choices				# obs.	Sign. of Conlisk's Z-statistic	Sign. of χ^2 test
Category		AB	A*B*	AB*	A*B			
Gender	Female	5.7	66.7	20.7	6.9	246	***	n.s.
	Male	2.9	73.4	16.5	7.2	278	***	
Age	Age 16–24	0.0	66.7	22.2	11.1	36	n.s.	n.s.
	Age 25–34	1.0	77.7	16.5	4.9	103	***	
	Age 35–44	4.6	73.6	12.6	9.2	87	n.s.	
	Age 45–54	4.9	64.8	23.8	6.6	122	***	
	Age 55–64	3.3	70.0	17.8	8.9	90	**	
	Age 65+	9.3	67.4	18.6	4.7	86	***	
Education	Primary education	3.3	66.7	16.7	13.3	30	n.s.	***
	Lower secondary education	9.1	59.8	24.2	6.8	132	***	
	Higher secondary education	2.8	73.6	15.3	8.3	72	n.s.	
	Intermediate vocational training	2.8	63.2	23.6	10.4	106	***	
	Higher vocational training	3.3	76.4	14.6	5.7	123	**	
	University degree	0.0	90.2	9.8	0.0	61	***	
Occupation	Employed (contract)	4.0	73.7	16.4	5.8	27	***	n.s.
	Freelance or self-employed	0.0	70.0	25.0	5.0	20	**	
	Unemployed	0.0	50.0	30.0	20.0	10	n.s.	
	Student	0.0	71.4	20.0	8.6	35	n.s.	
	Works in own household	9.5	58.7	20.6	11.1	63	*	
	Retired	4.5	75.0	15.9	4.5	88	***	
	Other occupation	2.9	55.9	29.4	11.8	34	*	

Table 8 (continued)

Treatment LowRᴇᴀʟ Category	Relative freq. of choices				# obs.	Sign. of Conlisk's Z-statistic	Sign. of χ² test
	AB	A*B*	AB*	A*B			
Household	5.5	56.8	24.7	13.0	146	***	***
Income							
HH gross income ≤ €2250	3.3	74.8	17.1	4.9	123	***	
HH gross income €2251–€3130	3.7	77.0	13.3	5.9	135	**	
HH gross income €3131–€4350	4.2	74.2	18.3	3.3	120	***	
HH gross income ≥ €4351							
Holds assets							*
Yes	3.7	81.5	9.9	4.9	443	***	
No	4.3	68.2	20.1	7.4	81	n.s.	
Has savings							n.s.
Yes	3.4	70.6	18.5	7.5	259	***	
Account							
No	5.0	69.9	18.5	6.6	265	***	

Notes: LowReal stands for low real payoffs. The column labeled "Sign. of Conlisk's Z-statistic" indicates significance for the test that violations of EUT are systematic (more AB^* than A^*B violations). The column labeled "Sign. of χ² test" indicates significance levels of χ² tests for differences between proportions of choices in the category listed in column 1. ***, **, * indicates significance at the 1%, 5% and 10% level, respectively; n.s. indicates non-significance

Table 9 Test results for violations of EUT across treatments

Category		Significance of Conlisk's D-statistic	
		HIGHHYP VS LOWHYP	LOWHYP VS LOWREAL
Gender	Female	*n.s.*	*n.s.*
	Male	***	**
Age	Age 16–24	**	*n.s.*
	Age 25–34	*	*n.s.*
	Age 35–44	***	*n.s.*
	Age 45–54	***	**
	Age 55–64	**	*n.s.*
	Age 65+	***	*n.s.*
Education	Primary education	*n.s.*	*
	Lower secondary education	***	*n.s.*
	Higher secondary education	***	*n.s.*
	Intermediate vocational training	***	**
	Higher vocational training	***	*n.s.*
	University degree	***	*n.s.*
Occupation	Employed (contract)	***	***
	Freelance or self-employed	*n.s.*	*
	Unemployed	*	*n.s.*
	Student	**	*n.s.*
	Works in own household	*n.s.*	*n.s.*
	Retired	***	*n.s.*
	Other occupation	***	*n.s.*
Household	HH gross income $\leq$ €2250	***	*
Income	HH gross income €2251−€3130	**	*n.s.*
	HH gross income €3131−€4350	***	*n.s.*
	HH gross income $\geq$ €4351	**	*
Holds assets	Yes	***	*n.s.*
	No	***	**
Has savings	Yes	***	*n.s.*
Account	No	***	*

Notes: HighHyp stands for high hypothetical payoffs, LowHyp for low hypothetical payoffs, and LowReal for low real payoffs. The columns labeled "Sign. of Conlisk's D-statistic" indicate significance for the test that violations of EUT are statistically significantly different across the two treatments listed in columns 3 and 4. ***, **, * indicates significance at the 1%, 5% and 10% level, respectively; *n.s.* indicates non-significance

Continue

Screen 9:
This is the end of the questionnaire. Thank you for your participation.

Appendix B: Relative frequencies of choices depending on socioeconomic characteristics

In Tables 5–9 we report the relative frequency of choices and the re-
sults of additional tests depending on socioeconomic characteristics. We
do this for the pooled data (Table 5) and the three treatments separately
(Tables 5–8), and for pair-wise across-treatment tests (Table 9). The various
tests are described in the notes to the tables.

Appendix C: Results of multinomial logit regressions

In this appendix, we report the results of multinomial logit regressions on all four answers AB, A^*B^*, AB^*, and A^*B, using expected value maximization, A^*B^*, as the base outcome and using the variables listed in column 1 in Table 3 as regressors. We perform multinomial logit regressions for the pooled data and for the three treatments separately. The results are reported in Tables 10 (all data) and Tables 11–13 (treatments HighHyp, LowHyp, and LowReal). In these tables we report the relative risks of choosing outcome AB, AB^* or A^*B over the base outcome A^*B^*. That is, the three columns in Tables 10–13 show, respectively, the ratios $P(answer = AB)/P(answer = A^*B^*)$, $P(answer = A^*B)/P(answer = A^*B^*)$, and $P(answer = AB^*)/P(answer = A^*B^*)$, where $P(.)$ denotes the probability of choosing a given answer. The tables should be read as follows. Refer, for example, to Table 10 that reports the results

Table 10 Results of a multinomial logistic regression (all treatments)

All treatments	Answer					
Base outcome: A^*B^*	AB		AB^*		A^*B	
Constant	0.189*	(−1.85)	0.547	(−0.99)	0.839	(−0.24)
HighHyp	14.116***	(9.57)	6.209***	(9.91)	8.665***	(8.19)
LowReal	1.043	(0.13)	1.309	(1.55)	1.398	(1.25)
Female	2.045***	(3.00)	1.297	(1.64)	0.929	(−0.32)
Age 25–34	0.726	(−0.44)	0.986	(−0.03)	0.175*	(−1.75)
Age 35–44	0.654	(−0.57)	0.810	(−0.42)	0.197***	(−2.73)
Age 45–54	0.877	(−0.18)	1.148	(0.28)	0.363*	(−1.77)
Age 55–64	1.381	(0.43)	1.201	(0.36)	0.297**	(−2.01)
Age 65+	1.441	(0.44)	0.895	(−0.20)	0.173**	(−2.49)
Lower second. edu.	0.571	(−1.28)	0.716	(−1.03)	0.927	(−0.17)
Higher second. edu.	0.449	(−1.64)	0.487**	(−2.03)	0.746	(−0.62)
Intermed. voc. training	0.418*	(−1.86)	0.784	(−0.72)	0.788	(−0.51)
Higher voc. training	0.260***	(−2.77)	0.683	(−1.12)	0.716	(−0.71)
University degree	0.175***	(−2.95)	0.475*	(−1.96)	0.093***	(−2.85)
Employed (contract)	0.811	(−0.51)	0.479***	(−2.59)	0.538	(−1.54)
Freelance or self-empl.	1.319	(0.46)	0.441*	(−1.72)	0.282	(−1.51)
Unemployed	0.000	(−0.02)	0.871	(−0.26)	1.437	(0.55)
Student	1.196*	(−1.77)	0.622	(−0.88)	0.259**	(−1.99)
Works in own household	0.861	(−0.33)	0.645	(−1.34)	0.808	(−0.46)
Retired	0.396*	(−1.68)	0.711	(−0.91)	1.280	(0.47)
HH gr. inc. €2251–€3130	0.966	(−0.12)	0.820	(−1.00)	0.432***	(−2.95)
HH gr. inc. €3131–€4350	0.670	(−1.32)	0.534***	(−3.00)	0.507**	(−2.50)
HH gr. inc. ≥ €4351	0.566*	(−1.71)	0.728	(−1.50)	0.349***	(−3.22)
Assets	1.032	(0.10)	0.666*	(−1.90)	0.578*	(−1.65)
Savings account	0.803	(−0.98)	1.306*	(1.79)	1.187	(0.81)
No. of observations	1424					
Log likelihood	−1321.24					

Relative risk ratios are reported

Notes: HighHyp stands for high hypothetical payoffs, LowHyp for low hypothetical payoffs, and LowReal for low real payoffs. Absolute value of z-statistics in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Omitted categories are treatment LowHyp; age interval [16–24]; primary education; "other" occupation; household gross income smaller or equal to €2250

Table 11 Results of a multinomial logistic regression (Treatment HιGHHγP)

Treatment HιGHHγP	Answer					
Base outcome: A^*B^*	AB		AB^*		A^*B	
Constant	4.663	(1.069)	3.059	(0.875)	2.670	(0.589)
Female	1.997**	(2.003)	0.991	(−0.031)	0.992	(−0.022)
Age 25–34	0.324	(−0.956)	1.408	(0.324)	0.218	(−1.334)
Age 35–44	0.129*	(−1.736)	0.542	(−0.580)	0.045***	(−2.657)
Age 45–54	0.228	(−1.253)	0.835	(−0.171)	0.139*	(−1.739)
Age 55–64	0.382	(−0.801)	0.905	(−0.090)	0.075**	(−2.146)
Age 65+	0.134	(−1.484)	0.675	(−0.322)	0.045**	(−2.283)
Lower second. edu.	0.287*	(−1.733)	0.719	(−0.475)	4.506	(1.278)
Higher second. edu.	0.390	(−1.229)	0.853	(−0.216)	3.460	(1.014)
Intermed. voc. training	0.306	(−1.580)	0.965	(−0.050)	2.731	(0.819)
Higher voc. training	0.276*	(−1.708)	1.484	(−0.552)	4.847	(1.303)
University degree	0.269	(−1.587)	1.249	(0.295)	1.029	(0.021)
Employed (contract)	1.441	(0.491)	0.417	(−1.466)	0.580	(−0.701)
Freelance or self-empl.	2.323	(0.887)	0.122**	(−2.085)	0.172	(−1.321)
Unemployed	0.000	(−0.025)	0.397	(−0.774)	0.878	(−0.099)
Student	0.186	(−1.215)	0.538	(−0.573)	0.114	(−1.635)
Works in own household	1.838	(0.752)	0.506	(−0.997)	0.589	(−0.594)
Retired	1.735	(0.599)	0.676	(−0.520)	2.349	(0.909)
HH gr. inc. €2251–€3130	1.766	(1.232)	1.389	(0.810)	0.646	(−0.882)
HH gr. inc. €3131–€4350	0.789	(−0.509)	0.465*	(−1.843)	0.429*	(−1.726)
HH gr. inc. ≥ €4351	0.617	(−0.959)	0.736	(−0.728)	0.412	(−1.631)
Assets	1.051	(0.115)	0.648	(−1.159)	0.620	(−0.940)
Savings account	0.898	(−0.321)	1.334	(0.991)	1.652	(1.378)
No. of observations	400					
Log likelihood	−490.83					

Relative risk ratios are reported

Notes: HighHyp stands for high hypothetical payoffs. Absolute value of z-statistics in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Omitted categories are treatment LowHyp; age interval [16–24]; primary education; "other" occupation; household gross income smaller or equal to €2250. One occupation observation missing

on the pooled data. Since the omitted treatment is LowHγp, the coefficient 14.116 in the second column of this table means that by moving from treatment LowHγp to treatment HιGHHγP, the relative risk, $P(answer = AB)/P(answer = A^*B^*)$, of choosing answer AB over answer A^*B^* is equal to 14.116. Similarly, by moving from the omitted age category 16–24 to age category 25–34, the relative risk, $P(answer = AB)/P(answer = A^*B^*)$, of choosing answer AB over answer A^*B^* is equal to 0.726.

Inspecting the results of the multinomial logit regressions for the three treatments in Tables 11–13, the most salient feature seems to be that for the different treatments different categories of background characteristics have a significant effect on the relative risk of choosing one answer over the base answer A^*B^*. For treatment HιGHHγP we note that it is most age categories that show a significant correlation with the relative risk of choosing answer

Table 12 Results of a multinomial logistic regression (Treatment LowHʏᴘ)

Treatment LowHʏᴘ	Answer					
Base outcome: A^*B^*	$A B$		$A B^*$		$A^* B$	
Constant	0.000	(−0.013)	0.793	(−0.222)	0.545	(−0.448)
Female	1.519	(0.764)	1.710*	(1.739)	0.926	(−0.157)
Age 25–34	9.0e+05	(0.013)	0.908	(−0.109)	2.132	(0.664)
Age 35–44	1.3e+06	(0.013)	1.128	(0.137)	0.256	(−0.902)
Age 45–54	6.1e+05	(0.012)	0.894	(−0.127)	1.249	(0.187)
Age 55–64	1.8e+06	(0.013)	1.520	(0.468)	1.091	(0.068)
Age 65+	2.6e+06	(0.014)	0.449	(−0.813)	0.358	(−0.717)
Lower second. edu.	0.771	(−0.288)	0.307**	(−2.271)	0.434	(−1.150)
Higher second. edu.	0.515	(−0.654)	0.203***	(−2.638)	0.436	(−1.049)
Intermed. voc. training	0.269	(−1.259)	0.307**	(−2.112)	0.439	(−1.022)
Higher voc. training	0.078*	(−1.883)	0.332*	(−1.959)	0.271	(−1.507)
University degree	0.000	(−0.014)	0.197**	(−2.347)	0.000	(−0.016)
Employed (contract)	0.257**	(−2.117)	0.353**	(−2.038)	0.445	(−1.024)
Freelance or self-empl.	0.348	(−0.852)	0.479	(−0.813)	0.000	(−0.008)
Unemployed	0.000	(−0.007)	0.891	(−0.139)	1.298	(0.193)
Student	0.000	(−0.011)	0.515	(−0.661)	1.093	(0.070)
Works in own household	0.102***	(−2.669)	0.749	(−0.507)	0.674	(−0.420)
Retired	0.031***	(−2.668)	1.449	(0.574)	2.733	(0.976)
HH gr. inc. €2251–€3130	0.687	(−0.640)	0.884	(−0.336)	0.460	(−1.249)
HH gr. inc. €3131–€4350	0.424	(−1.185)	0.742	(−0.776)	0.958	(−0.083)
HH gr. inc. ≥ €4351	0.276	(−1.479)	0.701	(−0.901)	0.311	(−1.609)
Assets	0.299	(−1.082)	1.011	(0.029)	0.603	(−0.619)
Savings account	0.897	(−0.214)	1.728*	(1.942)	0.607	(−1.171)
No. of observations	500					
Log likelihood	−346.80					

Relative risk ratios are reported

Notes: LowHyp stands for low hypothetical payoffs. Absolute value of z-statistics in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Omitted categories are treatment LowHʏᴘ; age interval [16–24]; primary education; "other" occupation; household gross income smaller or equal to €2250. One occupation observation missing

Table 13 Results of a multinomial logistic regression (Treatment LowRᴇᴀʟ)

Treatment LowRᴇᴀʟ	Answer					
Base outcome: A^*B^*	$A B$		$A B^*$		$A^* B$	
Constant	0.000	(−0.011)	0.922	(−0.075)	1.986	(0.493)
Female	2.012	(1.296)	1.227	(0.772)	0.707	(−0.820)
Age 25–34	7.5e+04	(0.007)	0.340	(−1.251)	0.120*	(−1.835)
Age 35–44	4.0e+05	(0.009)	0.325	(−1.235)	0.325	(−0.978)
Age 45–54	4.5e+05	(0.009)	0.684	(−0.431)	0.229	(−1.274)
Age 55–64	3.4e+05	(0.008)	0.462	(−0.850)	0.270	(−1.101)
Age 65+	3.6e+06	(0.010)	0.945	(−0.055)	0.241	(−1.003)
Lower second. edu.	2.383	(0.773)	1.910	(1.128)	0.784	(−0.347)
Higher second. edu.	0.724	(−0.247)	0.937	(−0.104)	0.825	(−0.256)
Intermed. voc. training	1.099	(0.073)	2.227	(1.323)	1.405	(0.464)
Higher voc. training	0.873	(−0.110)	1.131	(0.199)	0.816	(−0.267)

Table 13 (continued)

Treatment LowReal	Answer					
Base outcome: A^*B^*	AB		AB^*		A^*B	
University degree	0.000	(−0.011)	0.556	(−0.831)	0.000	(−0.014)
Employed (contract)	2.672	(0.830)	0.573	(−1.177)	0.576	(−0.801)
Freelance or self-empl.	0.000	(−0.006)	1.005	(0.007)	0.592	(−0.422)
Unemployed	0.000	(−0.003)	1.009	(0.010)	1.976	(0.630)
Student	0.000	(−0.008)	0.365	(−1.048)	0.215	(−1.153)
Works in own household	1.690	(0.435)	0.612	(−0.899)	1.212	(0.253)
Retired	0.277	(−0.955)	0.323	(−1.637)	0.391	(−0.911)
HH gr. inc. €2251–€3130	0.415	(−1.284)	0.542*	(−1.823)	0.250***	(−2.647)
HH gr. inc. €3131–€4350	0.360	(−1.553)	0.450**	(−2.312)	0.344**	(−2.216)
HH gr. inc. ≥ €4351	0.768	(−0.393)	0.764	(−0.781)	0.243**	(−2.313)
Assets	1.116	(0.153)	0.426**	(−1.993)	0.624	(−0.790)
Savings account	0.687	(−0.740)	1.083	(0.318)	1.207	(0.491)
No. of observations	524					
Log likelihood	−409.60					

Relative risk ratios are reported

Notes: LowReal stands for low real payoffs. Absolute value of z-statistics in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Omitted categories are treatment LowHyp; age interval [16–24]; primary education; "other" occupation; household gross income smaller or equal to €2250

A^*B over answer A^*B^* (columns labeled "A^*B" in Table 11). For treatment LowHyp we observe that most of the occupation variables are significantly correlated with the relative risk of choosing answer AB^* over answer A^*B^* (column labeled "AB^*" in Table 12). Finally, for treatment LowReal we infer that most of all the household gross income variables have a significant correlation with the relative risk of choosing answer A^*B over answer A^*B^* (column "A^*B" in Table 13).

References

Allais, M. (1953). Le comportement de l'homme rationnel devant le risqué: Critique des postulats et axioms de l'ecole americaine. *Econometrica, 21*, 503–546.

Allais, M., & Hagen, O. (Eds.) (1979). *Expected utility hypotheses and the Allais paradox*. Dordrecht: Reidel.

Andersen, S., Harrison, G. W., Lau, M. I., & Rutström, E. E. (2008). Risk aversion in game shows. In J. C. Cox, & G. W. Harrison (Eds.), *Risk aversion in experiments* (Vol. 12). Greenwich: JAI Press. Research in Experimental Economics.

Banks, J. (2006). Economic choices, capabilities and outcomes at older ages. *Fiscal Studies, 27*, 281–311.

Bellemare, C., & Kröger, S. (2007). On representative social capital. *European Economic Review, 51*, 183–202.

Benjamin, D. J., Brown, S. A., & Shapiro, J. M. (2006). Who is "behavioral"? Cognitive ability and anomalous preferences. Working paper.

Burke, M. S., Carter, J. C., Gominiak, R. D., & Ohl, D. F. (1996). An experimental note on the Allais paradox and monetary incentives. *Empirical Economics, 21*, 617–632.

Burks, S. V., Carpenter, J. P., Götte, L., & Rustichini, A. (2009). Cognitive skills affect economic preferences, strategic behavior, and job attachment. *Proceedings of the National Academy of Science, 106*, 7745–7750.

Camerer, C. (1995). Individual decision making. In J. H. Kagel, & A. E. Roth (Eds.), *The handbook of experimental economics* (pp. 587–703). Princeton: Princeton University Press.

Carbone, E. (2005). Demographics and behaviour. *Experimental Economics, 8*, 217–232.

Conlisk, J. (1989). Three variants on the Allais example. *American Economic Review, 79*, 392–407.

Dohmen, T., Falk, A., Huffman, D., & Sunde, U. (2010). Are risk aversion and impatience related to cognitive ability? *American Economic Review, 100*, 1238–1260.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., & Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants and behavioral consequences. *Journal of the European Economic Association, 9*, 522–550.

Eckel, C., & Grossman, P. (2000). Volunteers and pseudo-volunteers: The effect of recruitment method on subjects' behavior in experiments. *Experimental Economics, 3*, 107–120.

Fan, C.-P. (2002). Allais paradox in the small. *Journal of Economic Behavior and Organization, 49*, 411–421.

Fatas, E., Neugebauer, T., & Tamborero, P. (2007). How politicians make decisions: A political choice experiment. *Journal of Economics, 92*, 167–196.

Gächter, S., Huck, S., & Weizsäcker, G. (2008). Socio-demographics and choice in experimental economics. Mimeo.

Harrison, G. W. (1994). Expected utility and the experimentalists. *Empirical Economics, 19*, 223–253.

Harrison, G. W., & List, J. A. (2004). Field experiments. *Journal of Economic Literature, 42*, 1013–1059.

Harrison, G. W., Lau, M. I., & Rutström, E. E. (2009). Risk attitudes, randomization to treatment, and self-selection into experiments. *Journal of Economic Behavior and Organization, 70*, 498–507.

Heckman, J. (1976). The common structure of statistical models of truncation, sample selection, and limited dependent variables and a simple estimator for such models. *Annals of Economic and Social Measurement, 5*, 475–492.

Hey, J. D. (2002). Experimental economics and the theory of decision making under risk and uncertainty. *Geneva Papers on Risk and Insurance Theory, 27*, 5–21.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*, 263–291.

Levitt, S. D., & List, J. A. (2007). What do laboratory experiments measuring social preferences reveal about the real world. *Journal of Economic Perspectives, 21*, 153–174.

List, J. A., & Haigh, M. S. (2005). A simple test of expected utility theory using professional traders. *Proceedings of the National Academy of Science, 102*, 945–948.

MacCrimmon, K. R. (1968). Descriptive and normative implications of the decision-theory postulates. In K. Borch, & J. Mossin (Eds.), *Risk and uncertainty* (Chapter 1). London: Macmillan.

Machina, M. J. (1982). Expected utility analysis without the independence axiom. *Econometrica, 50*, 277–323.

Post, T., van den Assem, M. J., Baltussen, G., & Thaler, R. H. (2008). Deal or no deal? Decision making under risk in a large-payoff game show. *American Economic Review, 98*, 38–71.

Slovic, P., & Tversky, A. (1974). Who accepts Savage's axiom? *Behavioral Sciences, 19*, 368–373.

Starmer, C. (2000). Developments in non-expected utility theory: The hunt for a descriptive theory of choice under risk. *Journal of Economic Literature*, 38, 332–82.

Starmer, C., & Sugden, R. (1991). Does the random-lottery incentive system elicit true preferences? An experimental investigation. *American Economic Review, 81*, 971–978.

van de Kuilen, G., & Wakker, P. P. (2006). Learning in the Allais paradox. *Journal of Risk and Uncertainty, 13*, 155–164.

Viscusi, W. K. (1989). Prospective reference theory: Toward an explanation of the paradoxes. *Journal of Risk and Uncertainty, 2*, 235–264.

von Gaudecker, H.-M., van Soest, A., & Wengström, E. (2011a). Heterogeneity in risky choice behaviour in a broad population. *American Economic Review, 101*, 664–694.

von Gaudecker, H.-M., van Soest, A., & Wengström, E. (2011b). Experts in experiments: How selection matters for estimated distributions of risk preferences. IZA Discussion Paper No. 5575.