MASTER

ECONOMICS

MASTER'S FINAL WORK

DISSERTATION

MEASURING RISK APPETITE IN THE PORTUGUESE GENERATION Z: AN EXPERIMENTAL APPROACH

OLAF GHANIZADEH

SUPERVISION:

MARIA JOANA DANTAS VAZ PAIS RIBEIRO

OCTOBER - 2021

Abstract

We implement a Holt & Laury (2002) Multiple Price List experiment tested on a specific sample of the Portuguese population. By analyzing a sample of 90 Portuguese members of Generation Z, we find that subjects still make risk-averse choices even though we use lottery tickets as a reward. The findings are broadly consistent with previously published implementations of the experiment. We find that females are slightly more risk averse than males, but no statistically significant relationship in other measured factors such as self-reporting of risk attitudes, choice of investment product or knowledge of economics.

Journal of Economic Literature (JEL) Codes: C90; C91; D81; D90; D91

Contents

List of Figures

List of Tables

Acknowledgements

First of all, I wish to thank Professor Joana Pais for her invaluable assistance, patience, and availability in coordinating and performing this study and for opening doors and presenting me with opportunities that have allowed me to capitalize on my skills. Furthermore, I wish to thank Inês Dias, the lab manager at XLAB, for her availability to help me perform this study on such short notice.

Concluding a Master's degree and writing a dissertation is no small feat, and it would not have been possible without the support of my family and friends. In particular, I would like to thank my partner, Julia, for always being by my side and believing in me. My father, Ahmad, for always encouraging me when the goings were tough.

I would particularly like to thank my friends who, in some way or another, helped me in the process of concluding this dissertation. In no particular order; Luís Estevão Gonçalves García for continuous input and suggestions, Rui Santos for solving issues in LaTeXand Marius Larsson for helping me with data cleaning and analysis in R.

Furthermore, this work would not have been possible if it had not been for the maintainers, contributors and the communities of the free, open-source software used in this project. Without tools such as oTree, R, tidyverse and rmarkdown, this project would have been considerably harder to execute.

Measuring Risk appetite in the Portuguese Generation Z: An Experimental Approach

Olaf Ghanizadeh

1 Introduction

Decisions impacted by economic risk are a large part of modern life and have been a topic of interest for economists for centuries. Our preferences towards risk is to a large extent shaped by the society we live in. In a rapidly evolving world, it is natural to believe that societal changes that have occurred between generations may shape risk preferences, both for individuals and a population as a whole. Experimental economics allows us to test our theories using real subjects faced with incentivized decisions, giving proper motivation for subjects to answer truthfully. Although research have been made on Generation Z, there is not much literature utilizing experimental methodology to study the members of this generation and their preferences towards risk, even less so on the Portuguese Generation Z. This paper aims to contribute to the existing literature by applying the well tested Multiple Price List method from Holt & Laury (2002) to estimate the risk preferences of the Portuguese Generation Z. In addition, we want to determine whether important demographic and socio-economic indicators predict the behavior of subjects.

The Multiple Price List, hereafter MPL, is a well-known experimental method that has been used in a large variety of contexts to estimate risk preferences. This method asks subjects to choose between a series of paired lotteries. Each lottery consists of two possible outcomes: a higher and lower payoff outcome. Each outcome is assigned to a probability of that outcome being drawn at random if this lottery is chosen over the other. As the subjects move down the list of decisions, the probability of the higher payoff outcome being drawn increases, which is shown to subjects as a percentage representation of the probability.

In each decision, one lottery is relatively safer than the other, meaning that the variance between the payoffs in the two outcomes is smaller. As the probabilities in the lotteries changes, there is a point where the relatively safer lottery changes place with the riskier lottery. This is known as the switching point, which can be used to estimate a subject's willingness to take risks. If a subject makes a larger share of safe choices, we can infer their risk aversion is higher.

One of its main advantages is that it allows experimenters to use the expected utility theory (EUT) framework to analyse revealed preferences. The estimations the method provides are relatively precise within the above mentioned framework, and results are therefore relatively easy to interpret. The procedure is designed so that there are sufficient incentives for subjects to reveal their real preferences when choosing between a set of paired lotteries. Our implementation uses lottery tickets that gives subjects the possibility to win prizes from a basket of popular consumer electronics.

A specific characteristic of Generation Z (b. 1997 - 2012) is that they have had access to the Internet throughout their childhood and adolescence. A large part of this generation has grown to see the Internet take an increasingly more critical role in society. According to Statistics Portugal (2020), residential Internet access in Portugal has evolved from 15,1% in 2002 to 84,5% in 2020. As this generation is now entering the workforce, many will have started entering capital markets. This generation has lived through a severe global recession (the global financial crisis of 2008) in their adolescent years, which affected Portugal severely in the form of austerity programs. In addition, the access to Internet has made it much easier to enter capital markets, and the offering of riskier financial products, such as cryptocurrencies, binary options and other complex, often leveraged financial products, has widened. Considering the increased accessibility of these risky products and the reasons mentioned above, we want to find out if this has affected the risk preferences of the Portuguese Generation Z by

using well-tested experimental methodology.

We find that subjects make risk-averse choices even when lottery tickets are used as rewards. When investigating the relationship between risk aversion and gender, female participants are slightly more risk-averse than male participants. No statistically significant relationship is found in other factors such as self-reporting, knowledge of economics or choice of investment products.

2 Literature review

Over the years, several methods for risk preference elicitation using experimental methodology have emerged. However, the MPL method, first used in Binswanger (1980) and later popularized in Holt & Laury (2002), stands out as a framework that provides easily interpretable results while being relatively simple to replicate and apply to various situations with or without minor modifications. One of the main advantages of the method developed in Holt & Laury (2002) is that it allows us to analyse the results within the framework of expected utility theory (EUT). By assuming a particular functional form over utility, the method allows experimenters to estimate rather precise intervals of a risk aversion parameter that can be used in further analysis.

It has grown to become a go-to approach for estimating risk preferences in both field and lab experiments. Filippin & Crosetto (2016) found in a survey that as of January 2013, 118 articles had been published replicating the original experiment (Holt & Laury, 2002) in some form or another.

The method used in Holt & Laury (2002) builds upon the results from Binswanger (1980), where a sample of farmers from rural India participated in both a questionnaire and an experiment where they were presented with a series of gambles consisting of coin tosses, finding that results were inconsistent across elicitation methods and that the degree of risk aversion tends to increase as payoffs are scaled up. Holt & Laury (2002) expands on this by using treatments that apply multipliers to the lottery prizes; in addition, some treatments involved hypothetical winnings.

Harrison et al. (2005) argues that the original experiment (Holt & Laury, 2002) confounds order and treatment effects by letting subjects first play a low payoff task, then a higher payoff treatment task. The authors argue that changes in behavior from scaling up the payoffs in consequent rounds result from both the order and scale effect. They run an experiment where subjects in one treatment are given a low payoff task followed by a higher payoff task (as in Holt & Laury, 2002) and a second treatment where subjects only face a high payoff task to isolate the order effect and the scale effect on behavior. They find that scaling payoffs has a much smaller effect on behavior than what is argued by Holt & Laury (2002). One of the main findings of Holt & Laury (2002) is that due to the increase in risk aversion when scaling up payoffs, the assumptions of CRRA utility may not hold. The findings of Harrison et al. (2005) argue that when the order effects are removed and the perceived scale effect is reduced, CRRA utility still holds.

In a follow-up study, Holt & Laury (2005) performs a new experiment aiming to test the effect of controlling for order effects. Although they agree with Harrison et al. (2005) that order effects are important, their data show that the scale effect is significant even when controlling for order effects, therefore reaffirming the main conclusions of Holt & Laury (2002).

Despite the fact that the MPL is easy to implement and results are easy to interpret for experimenters, it may be harder to understand for participants than other experimental methods for risk preference elicitation. This lack of comprehension may have several consequences, for example, subjects making inconsistent choices or so-called multiple switching behavior, hereafter MSB. In order to estimate the risk aversion parameter, one requires that subjects switch between the two lotteries once and retain this switching point. As subjects are free to decide between the two lotteries in each row, some subjects switch more than once, implying that their revealed preferences are not convex. Although MSB is the main form of inconsistent decision making, it may also manifest as subjects making their first choices at the bottom of the task list, or always picking the safe option, losing out on a certain, larger payoff in the last task.

Inconsistent behavior is expected in experiments using the MPL method and depending on how one chooses to address it. It may answer a different question, lower the overall data quality or introduce bias on the sample analysed.

Filippin & Crosetto (2016) surveyed 54 published papers replicating the HL method. The authors find that across a total sample of 6315 subjects, the inconsistency rate was on average 17,1% (Crosetto & Filippin, 2016). The seminal paper reports that 28 out of 212 (13,2%)

subjects displayed such behavior in the first round (low real payoffs) (Holt & Laury, 2002, p. 1647).

Given the common occurrence of inconsistent behavior in MPL experiments, a large body of literature has emerged on the topic. Different strategies applied to data analysis, different variations of the MPL, and other strategies are among the solutions proposed in the body as mentioned above of literature. The inconsistency rate seems to depend largely on the pool of participants; a survey of experimental methods for risk preference elicitation by Charness et al. (2013) recounts a study made on Rwandan adults where the inconsistency rate was at 55% and another of rural Senegalese farmers where the rate of inconsistency was 75%. Dave et al. (2010) found an inconsistency rate of 8.5% in a sample of 881 Canadian residents. These results support that the inconsistency rates vary across samples and that it is an important metric for experimenters to analyse. Andersen et al. (2006) argues that inconsistency rates may be a result of indifference between the two lotteries, but without giving the participants an option explicitly stating that they are indifferent between the two lotteries.

Variations of the MPL has been suggested in order to minimize or eliminate the risk of inconsistent behavior. One of these variants is the switching MPL (sMPL) or iterative MPL (iMPL) proposed by Andersen et al. (2006), where subjects are asked at which row they would prefer to switch and are not free to decide each row independently. Harrison et al. (2007) utilizes a variant of the iMPL on a sample of 253 Danish citizens, where the subjects are asked to indicate a point where they would like to switch, while also giving an indifference option, the rest of the choice sequence is then taken as given, assuming monotonicity of the subject's preferences. As noted by Yu et al. (2021), forcing a subject to state their preferred switching point eliminates the risk of MSB. However, it also reduces the choice set. One does not know if the subjects would display MSB if they had the opportunity to select each task independently. Yu et al. (2021) develops an extension to the MPL with a nudge protocol that allows the subjects to redo their choices in a given round if they felt insecure about their decision making. Ihli et al. (2016) compares inconsistency rates across two risk preference elicitation methods on a sample of rural Ugandans. In order to aid comprehension, the choices are presented visually to participants in both methods and report lower inconsistency rates than other comparable samples.

One way to deal with inconsistent participants is to drop those observations from the dataset. However, caution should be taken as this may introduce a selection bias.

Another option is to use the number of safe choices as a proxy for the switching point

Engel & Kirchkamp (2019) develops a framework that allows experimenters to use data on inconsistent participants to strengthen their parameter estimation results when using the MPL in conjunction with another risk preference elicitation task.

As argued by Charness et. al. (2013, 2020), even though a vast array of experimental risk preference elicitation methods exist, they have limited ability to predict real-world economic behavior. Policymakers should be careful with extrapolating the results from experimental studies into policy design.

3 Experimental design and implementation

The experiment utilizes the methodology devised by Holt & Laury (2002) and was implemented in oTree[1](Chen et al., 2016), which is an open source, Python-based, software for creating experiments. Our experiment stays true to the HL method, with a couple of modifications, one being that subjects are offered lottery tickets instead of monetary prizes due to funding constraints.

In the HL experiment, subjects face a series of ten choices between two paired lotteries. The participant can choose between lottery A or B in each choice, where A is the relatively safer option, and B is the relative riskier option. In the safe option, the variability between the possible outcomes is lower than in the riskier option. In our experiment the participants were asked to choose between the safer lottery, A, with possible winnings of 20 or 16 lottery tickets,

[1]The source code of the oTree implementation of experiment and its parts can be found at the author's GitHub: https://github.com/olafghanizadeh/oTree. A live version of the experiment and instructions can be found at: http://genz-welcome.herokuapp.com/

Table I: Payoff matrix of the HL lottery.

Task	Option A	Option B	$EV[A]$	$EV[B]$	Difference
1	With 0.1 win 20ℓ, with 0.9 win 16ℓ	With 0.1 win 38ℓ, with 0.9 win 1ℓ	16.4	4.7	13.5
2	With 0.2 win 20ℓ, with 0.8 win 16ℓ	With 0.2 win 38ℓ, with 0.8 win 1ℓ	16.8	8.4	10
3	With 0.3 win 20ℓ, with 0.7 win 16ℓ	With 0.3 win 38ℓ, with 0.7 win 1ℓ	17.2	12.1	6.5
4	With 0.4 win 20ℓ, with 0.6 win 16ℓ	With 0.4 win 38ℓ, with 0.6 win 1ℓ	17.6	15.8	3
5	With 0.5 win 20ℓ, with 0.5 win 16ℓ	With 0.5 win 38ℓ, with 0.5 win 1ℓ	18	19.5	-0.5
6	With 0.6 win 20ℓ, with 0.4 win 16ℓ	With 0.6 win 38ℓ, with 0.4 win 1ℓ	18.4	23.2	-4
7	With 0.7 win 20ℓ, with 0.3 win 16ℓ	With 0.7 win 38ℓ, with 0.3 win 1ℓ	18.8	26.9	-7.5
8	With 0.8 win 20ℓ, with 0.2 win 16ℓ	With 0.8 win 38ℓ, with 0.2 win 1ℓ	19.2	30.6	-11
9	With 0.9 win 20ℓ, with 0.1 win 16ℓ	With 0.9 win 38ℓ, with 0.1 win 1ℓ	19.6	34.3	-14.5
10	With 1 win 20ℓ, with 0 win 16ℓ	With 1 win 38ℓ, with 0 win 1ℓ	20	38	-18

or the riskier lottery where the possible outcomes were 38 or 1 lottery tickets. The number of lottery tickets are the same as the dollar amounts used in Holt & Laury (2002), scaled up by a factor of ten and rounded down to the closest integer. Along with each choice, a probability of each outcome is given, which changes as the participant moves down the list of choices. Denoting the probability of the the first outcome being drawn in any choice as p_i, and $1 - p_i$ for the second outcome being drawn, where p changes according to $p \in \{.1, .2, .3...1\}$ as participants move down the list of choices.

Table I shows each choice, the expected value of each option, and the difference between choosing the safe alternative over the riskier alternative, note that the expected values and difference is not shown to subjects during the experiment. From the table we see that a risk neutral individual would switch from option A to option B after the 4th task to maximize the expected value in each choice. A participant that is risk averse would pick option A in a larger share of the tasks than someone who is risk neutral, and a risk loving individual would pick option B in most of their choices.

The 10th choice serves as a test to check if the subject has understood the experiment. If the participant understood the experiment, they should always choose option B, yielding a certain payoff of 38 lottery tickets.

We can then use the switching point, or the number of safe choices, to estimate the risk aversion parameter.

3.1 CRRA Intervals and classification of risk aversion

As in Holt & Laury (2002) we use a utility function displaying constant relative risk aversion (CRRA), which takes the following form:

$$U(\ell) = \frac{\ell^{1-r}}{1-r} \tag{1}$$

Where U is the utility of lottery tickets, ℓ, and r is the Arrow–Pratt measure of relative risk aversion.

$$EU_i = \sum_{i=1}^{10}(p_i \times U_i) \tag{2}$$

Equation (2) is the expected utility of any choice in the list of tasks.

Table II summarizes the r values and classification levels for different number of safe choices. The intervals of r are determined by solving for r as show in (3) and (4) below. In this example the subject picks option A four times before switching to option B. By solving (3) we find the lower bound of the CRRA interval, and the upper bound is found by solving for r in (4), giving a CRRA interval of $[-0.16, 0.13]$ which translates to the risk neutral (RN) classification in Table II.

$$0.4\frac{20^{1-r}}{1-r} + 0.6\frac{16^{1-r}}{1-r} = 0.4\frac{38^{1-r}}{1-r} + 0.6\frac{1^{1-r}}{1-r} \Leftrightarrow r \equiv -0.16 \tag{3}$$

$$0.5\frac{20^{1-r}}{1-r} + 0.5\frac{16^{1-r}}{1-r} = 0.5\frac{38^{1-r}}{1-r} + 0.5\frac{1^{1-r}}{1-r} \Leftrightarrow r \equiv 0.13 \tag{4}$$

Table II: Risk aversion classification

Number of safe choices	CRRA	Classification
0-1	$r <$ -0.98	highly RL
2	-0.98 $< r <$ -0.51	very RL
3	-0.51 $< r <$ -0.16	RL
4	-0.16 $< r <$ 0.13	RN
5	0.13 $< r <$ 0.4	slightly RA
6	0.4 $< r <$ 0.66	RA
7	0.66 $< r <$ 0.96	very RA
8	0.96 $< r <$ 1.36	highly RA
9-10	1.36 $< r$	extremely RA

Generally speaking, $r = 0$ implies risk neutrality, $r < 0$ implies risk aversion and $r > 0$ implies risk loving preferences.

3.2 Incentive compatibility

An important modification in our experimental design, when comparing to HL, is the use of lottery tickets instead of monetary prizes. Economists generally posit that monetary prizes are necessary for truthful revalation of preferences as subjects do not necessarily have an intrinsic valuation of lottery tickets.[2]

However, Kirchkamp et al. (2021) finds that changing monetary prizes for tokens does not lead participants to make more risk neutral choices by comparing data across several sessions and previous studies using monetary prizes only. There, the exchange rate between tokens and euro is given to participants as a part of the instructions.

In this experiment, the lottery tickets give the participants the ability to participate in a final

[2]Even though some psychologists have argued that subjects have enough internal motivation to reveal truthful preferences, even in the absence of financial rewards (Camerer & Hogarth, 1999).

Table III: Treatments used in the experiment

Treatment	Round 1	Round 2	Round 3	Round 4
90HypReal	1x	90x	90x	1x
50HypReal	1x	50x	50x	1x
20Hyp	1x	20x	1x	-
20Real	1x	20x	1x	-

lottery with a prize pool consisting of an array of popular consumer electronics. This information along with the monetary value of the prize pool is given to the participants as a part of the instructions before starting the experiment. We believe this to be sufficient information for participants to reveal their real preferences in the the lottery choice tasks.

3.3 Treatments

In order to test whether higher amounts induces risk aversion in subjects we are utilizing five treatments, following the same structure as in Holt & Laury (2002). All participants play the baseline in the first round, meaning no multiplier and real rewards. This is equal to the Low Real Payoffs treatment from Holt & Laury (2002). From the second round and out the multiplier varies depending on the randomly assigned treatment. Three treatments include both a hypothetical round and a real round with the same multiplier, and the two remaining treatments consist of a single round with a 20x multiplier, where the reward is either hypothetical or real. After playing the treatment rounds, the subjects will play the baseline round once more. After each round the subjects are given the option to withdraw and leave the experiment with whatever winnings they have collected up until that point, this option is given regardless if the subjects played a round with hypothetical or real rewards. This is a modification with respect to Holt & Laury (2002) where subjects only were given the option to withdraw after the first round.

Table III summarizes the treatments used in the experiment.

3.4 Questionnaire

After completing the HL task, all participants were asked to complete a brief questionnaire to collect anonymous demographic and socioeconomic data about the participants. This data will be collected in order for us to be able to analyse subject behavior across a set of demographic variables, such as gender and education.

We added a question that asked participants to indicate whether they had previously studied economics, at least at an undergraduate level. This question was added to have a data point where we could analyse if participants who had previously studied economics to a more significant extent would maximize expected values in the lottery task and converge towards risk neutrality compared with a sample that does not indicate to have studied economics previously.

A question asks participants to self-evaluate their risk appetites was added towards the end of the questionnaire. The aim of this question was to see to what extent the preferences revealed in the lottery task correlated with the self-evaluation. Here we also added a question asking participants to indicate whether they had previously done any online gambling.

The final question in the questionnaire asks subjects to indicate which, if any, financial assets they own or have previously owned. The basket of assets was replicated from Portuguese Securities Market Commission's (CMVM) classification of financial market assets. We added this question to look at the correlation between revealed risk preferences and financial market asset preferences.

3.5 Experimental procedures

In total four sessions were held at the Lisbon School of Economics and Management (ISEG) experimental laboratory, XLAB,[3] over the course of two days in September 2021. Participants were mainly recruited through XLAB's recruitment system (ORSEE), first year Master students, and incoming undergraduate students. In total, 95 participants completed the ex-

[3]http://xlabonline.com/

periment. However, some participants who came to the lab to participate fell outside the recruitment criteria, leading to 5 observations being dropped[4] and turning the qualifying sample size to 90 . Our primary recruitment criterium required participants to be Portuguese, and members of Generation Z (b. 1997-2003) . For ethical reasons, the minimum age of a participant at XLAB is 18 years old, thus the last year of birth we could include in our sample was 2003 .

Subjects were randomly assigned a treatment upon starting the experiment to ensure a balanced distribution of subjects across treatments. Upon completing the experiment, the subjects were shown a screen that gave them their final number of lottery tickets and instructions on how to redeem prizes in case they won.

The prizes were drawn at the end of October 2021 under the supervision of the lab manager. Winners were drawn randomly from a pool of the total number of real tickets awarded. Each participant entered the lottery with the number of tickets real tickets they gained in all non-hypothetical rounds of the experiment.

Winners were notified via E-mail, and the prizes were distributed to winners over the following week.

3.6 Data and hypotheses

The data we obtained from the experiment allowed us to analyze both individual behavior and the behavior of groups defined by different subject characteristics obtained from survey questions. For each participant, our data shows which lottery each participant picked in each selected in each decision throughout the different rounds, which treatment they were assigned, whether they withdraw following that round and the number of times the switch between the safe and riskier lottery in each round. Combined with the questionnaire data this gave us the ability to test the following hypotheses.

[4]The dropped participants who completed were rewarded on the same basis as the qualifying participants

Hypothesis 1: Male subjects are less averse to risk One of the results presented in Holt & Laury (2002) is that when looking at demographic variables, is a slight difference between male and female subjects regarding the number of safe choices. We wanted to see whether this difference existed in our experiment as well. For this, we define the null and alternative hypotheses as follows:

$$H_0 : n_{\text{safe}_F} > n_{\text{safe}_M}$$

$$H_1 : n_{\text{safe}_F} = n_{\text{safe}_M}$$

Where n_{safe} is the number of safe choices for a sub-sample, in this case the male and female subjects. Our null hypothesis is that female participants makes more safe choices than the male subjects, and the alternative hypothesis is that there is no difference between them.

Hypothesis 2: Subjects with a background in economics are less averse to risk When studying economics, the concept of expected values is taught early on. In our experiment, a subject can make choices to maximize the expected value of each lottery. We wanted to test whether subjects with a background in economics would make such choices. In order to test this, we define the hypothesis that subjects with a background in economics would make fewer safe choices and opt to maximize the expected value of each lottery instead of making risk-averse choices. We can define this with a null and alternative hypothesis as follows:

$$H_0 : n_{\text{safe}_E} < n_{\text{safe}}$$

$$H_1 : n_{\text{safe}_E} = n_{\text{safe}}$$

Where n_{safe} is the sample of subjects that indicates that they have not studied economics previously. Therefore, our null hypothesis is that students with a background in economics make fewer safe choices to maximize expected values, and the alternative is that there is no difference between the two groups.

Hypothesis 3: Self-evaluation of risk does not correlate with behaviour In order to see whether subjects make choices that correlate with their stated perception of their willingness to take risks, we ask the subjects to indicate this on a scale that corresponds to the risk classification levels used to define the intervals in the experiment. We want to analyze the mean number of safe choices for each risk classification level. We can define the hypotheses as follows:

$$H_0 : \bar{n_{\text{safe}}}_{VRA} = \bar{n_{\text{safe}}}_{RA} = \bar{n_{\text{safe}}}_{RN} = \bar{n_{\text{safe}}}_{RL} = \bar{n_{\text{safe}}}_{VRL}$$

$$H_1 : \bar{n_{\text{safe}}}_{VRA} > \bar{n_{\text{safe}}}_{RA} > \bar{n_{\text{safe}}}_{RN} > \bar{n_{\text{safe}}}_{RL} > \bar{n_{\text{safe}}}_{VRL}$$

Where $\bar{n_{\text{safe}}}$ is the mean number of safe choices for a level of risk classification, the subscript is the group name abbreviated. Our null hypothesis is that there is no difference between the groups, while the alternative is that the mean number of safe choices corresponds to each level.

Hypothesis 4: Subjects who have gambled are less averse to risk Gambling is an activity that involves taking on risks. Suppose subjects indicate that they have previously gambled or used online gambling services. In that case, we expect this sample to display fewer safe choices than the rest of the population. We define the following hypotheses:

$$H_0 : n_{\text{safe}_G} < n_{\text{safe}}$$

$$H_1 : n_{\text{safe}_G} = n_{\text{safe}}$$

Where n_{safe_G} denotes the sample that indicates that they have gambled in the past, our null hypothesis is, therefore, that participants in the sample that indicates that they have gambled make fewer safe choices than the rest of the population. The alternative is that there is no difference.

4 Results

Our sample of 90 subjects consists of 57% female and 43% male participants. On average, the subjects were 20 years old (born in 2001), where the majority indicated that their main occupation was studying (88%). 56% answered that their highest level of completed education was upper secondary education, indicating that the majority of our sample was incoming undergraduate students. 53% indicated that they had studied economics, at least on an undergraduate level. The full summary statistics of our sample can be seen in Table VIII in the appendices.

Given that our sample consists mainly of students, and especially from a management and economics school, it is not representative of the Portuguese Generation Z population, and we must take caution before making conclusions about risk preferences for the population as a whole based on our limited sample. It is also worth noting that most of our participants were recruited in a very short amount of time in a limited number of sessions, we may have introduced a self-selection bias in our sample by only including those who were able to participate in our limited number of sessions.

4.1 Withdrawals

When monitoring the sessions we discovered that several participants chose to withdraw at one point or another during the experiment. The option to withdraw was given to the participants in between each round on the page where they could see their results from the previous round. In the original experiment from Holt & Laury (2002), the participants were given the option to withdraw after the first round of Low Real Payoffs. The participants were given the choice to continue and give up the winnings from the first round, or withdraw and leave with the current winnings. This is done in order to control for wealth effects, such that the wealth gained in the first round would not affect how they acted in the following treatment rounds with higher payoffs.

In our experiment, a total of participants 52 withdrew, 18 after the first round. The fact that our experimental implementation allowed participants to withdraw after each completed round does not explain why we saw such a high amount of withdrawals after the first round in contrast to Holt & Laury (2002) where there were no withdrawals.

The withdrawals, sorted by treatment and round, are displayed in Table IX in the appendix. Regardless of treatment, all participants played the baseline Low Real task in the first round. In addition, we saw several participants withdrawing after the second round, in a treatment that involved a hypothetical payoff, leading to a payoff of zero after completing the survey, even though the rounds involving hypothetical payoffs were clearly marked in each selection screen and after in the round summary screen.

The high amount of withdrawals weakens our data from the treatment rounds; therefore, our analysis in the following sections will focus on the first round. Furthermore, as in Holt & Laury (2002) we use the first round with low real rewards to examine the relationship between risk aversion and the other possible explanatory variables.

4.2 *Inconsistent choices*

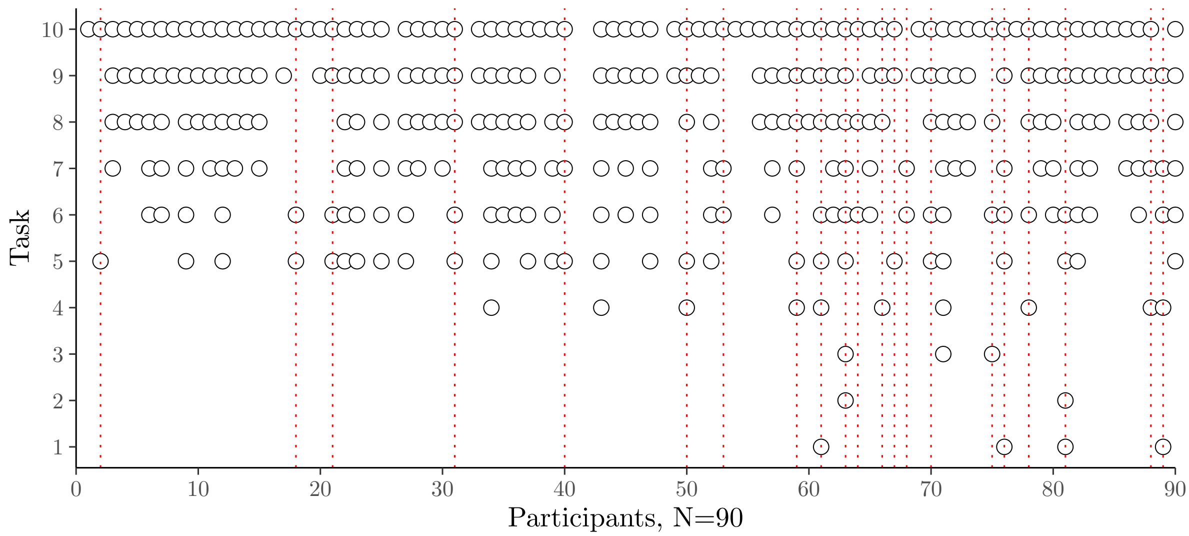

Figure 1: Inconsistent participants in the baseline round, adapted from Engel & Kirchkamp (2019). Each column is the choices of a participant, and each circle represents a risky choice. Participants marked with red lines are inconsistent, meaning that they haved displayed mutliple switching points.

In our baseline round, we saw that 22 out of 90 participants displayed MSB by switching more than once, meaning an overall MSB rate for the baseline round of 24,4%. As previously discussed, this inconsistent behavior may be an effect of the complexity of the HL design. We did not make any specific modifications to the design to minimize the risk of inconsistency other than translating all text in the experiment and instructions to Portuguese to reduce the risk of language barriers hindering task comprehension for subjects.

Figure 1 shows the distribution of risky choices (B) across all participants in the first baseline round. Inconsistent participants are marked with a red line. We can see that the bulk of inconsistent participants appears on the right side of the graph. This coincides with the sessions that were held mainly with incoming undergraduate students. Whether this is due to lack of comprehension or not taking the tasks seriously is not known.

Out of the 22 participants displaying MSB, 95% switched more than twice, most switching three times and two switching as much as seven times.

The choice sequence of a participant were (BAABABBABA) in the first round. The participant switched between the two lotteries seven times. This indicates that the participant is to some extent randomising between the two lotteries, notice especially that option A was chosen in the 10th task leading to a certain payoff of 20 instead of 38.

Although our inconsistency rate is above the average (Filippin & Crosetto, 2016), our sample is also relatively small. If we were to drop all inconsistent participants in further analysis, we may introduce a selection bias on the analysed sample, meaning that our results would be skewed favouring those who did not make inconsistent choices.

Furthermore, as in Holt & Laury (2002), we find very that there is only a small difference in the mean number of safe choices if we were to exclude the participants switching more than once. In our data set, by excluding the inconsistent behavior, we see the mean number of safe choices increase by 0.1. A two-sample Mann Withney U test of the number of safe choices in the full sample for the first round and a sample excluding MSB confirms ($p = 0.655, n_1 = 90, n_2 = 68$) that there is no statistically significant difference between the two samples.

As a result, we decided against dropping observations where participants switched back and forth between the two lotteries. However, a small discussion where we present the differences in the analysis of risk aversion when excluding the sample switching more than once is presented in the next section.

4.3 Risk aversion analysis

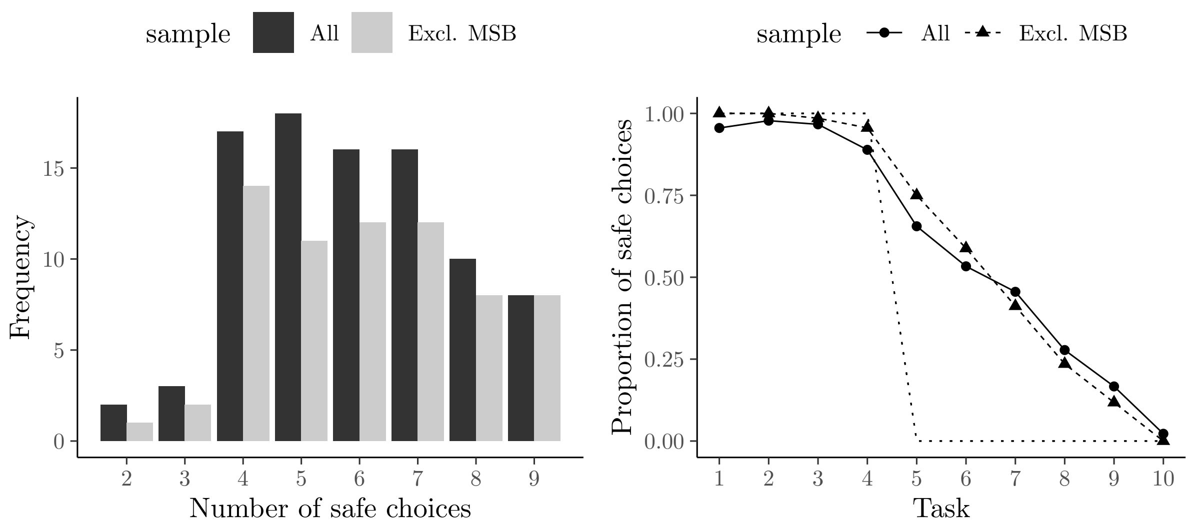

Figure 2: On the left, the frequency of safe choices in the first round. On the right, the proportion of safe choices per task, also in the first round, the dashed line indicates the risk neutral classification. The dashed line with triangles represents the sample excluding the subjects switching multiple times.

In order to estimate the risk aversion for our sample and differences between subjects and treatments, it is helpful to use both visual and numerical analysis. Figure 2 shows the aggregate of choices made by subjects in our sample in the baseline round. From the graphs, we can see that, even with lottery tickets, most subjects reveal risk-averse preferences. The graph to the left in illustrates the choices per task, and this graph can help us see the mean switching point between the two lotteries. The graph to the right shows that the frequency of safe choices is shifted to the right, meaning that most participants made four or more safe choices. The solid line with circles represents this. The solid line with triangles represents a sample where we exclude the participants displaying MSB. By excluding these subjects, the line takes a more uniform shape. As discussed previously, the difference between the two samples regarding the average number of safe choices is neither large nor statistically significant.

The dashed line indicates the perfectly risk-neutral classification, i.e. someone who maximizes the expected value in each choice. As the solid line is shifted to the right of the dashed line, it shows that on average the subjects made risk averse choices.

Table IV shows the risk classifications like in Table II with the proportions of choices of the samples including and excluding MSB to the right. Most subjects made five or more safe choices, further supporting the finding that most subjects reveal risk-averse preferences in the lottery choice task.

Table IV: Risk classification of subjects

			Proportion of subjects	
Number of safe choices	CRRA interval	Classifcation	All (n=90)	Excl. MSB (n=68)
0-1	$(-\infty$,-0.98]	highly RL	-	-
2	[-0.98,-0.51]	very RL	0.022	0.015
3	[-0.51,-0.16]	RL	0.033	0.029
4	[-0.16,0.13]	RN	0.189	0.206
5	[0.13,0.4]	slightly RA	0.200	0.162
6	[0.4,0.66]	RA	0.178	0.176
7	[0.66,0.96]	very RA	0.178	0.176
8	[0.96,1.36]	highly RA	0.111	0.118
9-10	[1.36 ,$+\infty$)	extremely RA	0.089	0.118

4.4 Gender differences in risk aversion

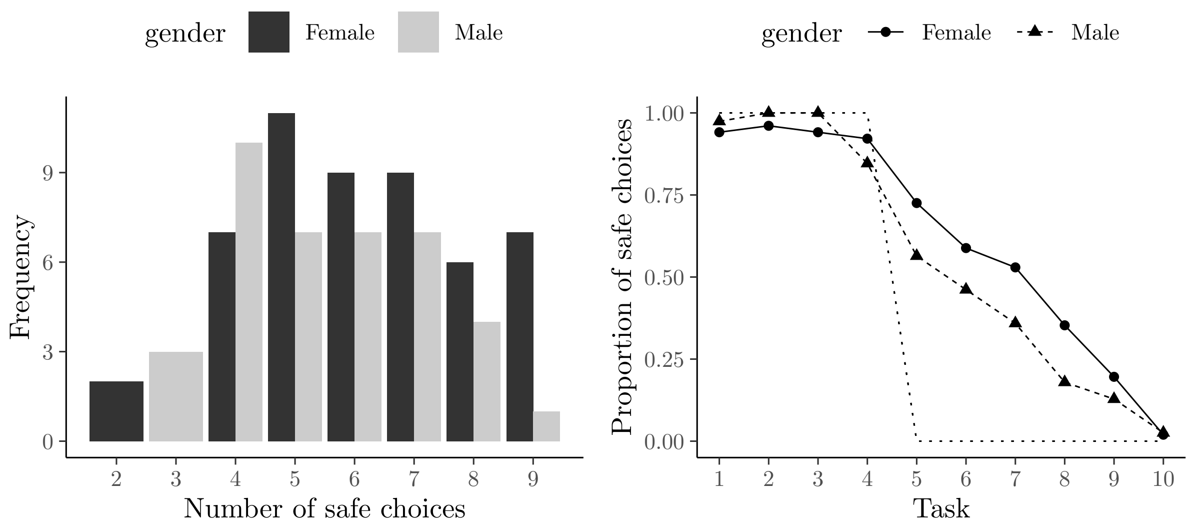

Figure 3: Safe choices across genders in the first baseline round

We used our data set and separated the choices in the baseline round based on gender. Doing the same visual analysis as above (Figure 3), we find that on average, male participants are less risk averse than female participants, where male participants on average made 5.53 safe choices, female participants made on average 6.2 safe choices in the baseline round. Although there exists a difference in the mean number of safe choice between the genders, when testing Hypothesis 1, the difference in safe choices between the genders is only significant at the 90% ($\alpha = 0.1$) confidence interval as measured by a Mann Whitney U Test ($p = 0.078$). Therefore, if we want to maintain a 5% significance level, we reject the null hypothesis and conclude that the difference is not statistically significant.

A similar finding is reported in Holt & Laury (2002) during the lower payoff rounds. In treatments with higher payoff multiplier there is no gender difference. While other studies report no gender difference in risk aversion when using the MPL method (Andersen et al., 2006; Harrison et al., 2007).

4.5 Are students with a background in economics less risk averse?

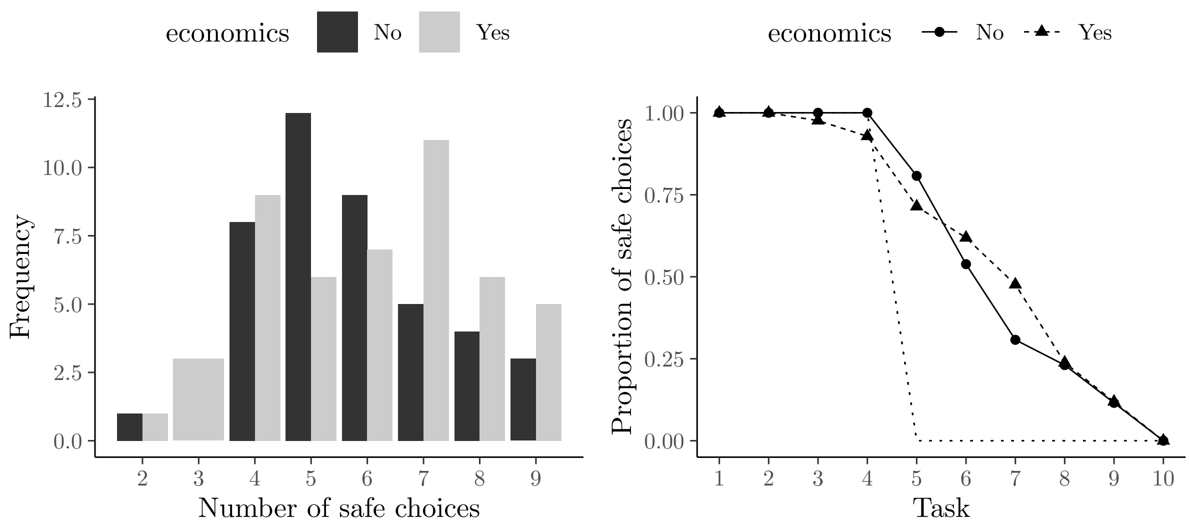

Figure 4: Safe choices between students who have a background in economics, in the first baseline round

The tests and trials of our experimental implementation were done mainly on students with a background in economics. It became apparent that many started calculating expected values for each task and used that as a decision rule. The concept of expected values is prevalent in many disciplines. However, students who have done courses in econometrics and microeconomics should be very familiar with this concept, and we thought this would lead to these students making choices that to a larger extent maximized the expected value of every choice. Therefore, we asked subjects to indicate in the questionnaire whether they had previously studied economics, at least at an undergraduate level.

We then use this data point to analyse whether these subjects maximize the expected values in every choice, even when there are real incentives involved. Figure 4 summarizes these results. From the graph above, it does seems like this sample, on average, initially converges towards the risk-neutral switching point. However, this effect is later leveled out by a larger share of safe choices in a later task. Whether this effect is caused by inconsistent behavior is not taken into account here, as we do not exclude subjects displaying such behavior.

We use a Mann Whitney U-Test to test Hypothesis 2, which gives $p = 0.46$. Therefore, we reject the null hypothesis and conclude that there is no statistically significant difference between the two samples. A similar result is reached in Holt & Laury (2002), where no significant difference in risk aversion was found when analyzing a series of demographic and socioeconomic variables, such as age, education level and field of studies.

4.6 Is self-evaluation of risk preferences predictive of behavior?

Table V: Mean number of safe choices per self-reported risk classification

Classifcation	n	Mean	Median	Proportion
Very risk averse	1	8.00	8	0.01
Risk averse	35	5.69	6	0.39
Risk neutral	23	6.35	6	0.26
Risk loving	28	5.82	5	0.31
Very risk loving	3	5.00	5	0.03

As a part of our questionnaire, we asked the subjects to indicate their perceived risk preferences on a scale similar to the classifications used in Table II. We wanted to see to what extent the choices made in the lottery task correlated with how subjects self-evaluated their risk preferences. Table V shows the mean number of safe choices categorized by the self-evaluated risk classification. By comparing with the numbers in Table II, we see that regardless of the subject's self-evaluation, the average subject made risk-averse choices in the lottery task, notice especially that both the risk loving classifications are well on the side of risk aversion according to Table II.

To test Hypothesis 3 we use a one-way ANOVA test, which allows us test whether there is a statistically significant difference between the means of our five groups. We find that the differences between the groups are not statistically significant ($p = 0.369$), and we fail to reject the null hypothesis. This result indicates that self-evaluation of risk is not predictive of actual behavior when placed in a situation involving incentivized decisions.

Similar results are reached in Dohmen et al. (2011), where simply asking for a subject's willingness to take risks have very little predictive power. However, when contextualizing the questions with different framing the predictive power of self-reports is stronger.

4.7 Does gambling affect risk aversion?

Table VI: Mean number of safe choices in subjects who indicated that they have used gambling services.

Classifcation	n	Mean	Median	Proportion
Non-gambler	62	5.95	6	0.69
Gambler	28	5.79	6	0.31

We asked the subjects to indicate if they had previous experience with online gambling services that allows you to place wagers on sports events. These services typically require you to use real money to bet on the outcome of certain events. This kind of behavior is often classified as risk-seeking. We were interested in seeing if subjects who indicated that they had used these services made fewer safe choices when compared to the group that does not have gambling experience.

Table VI shows summary statistics of the two samples in the first round of our experiment. The table shows that, on average, the participants who indicated that they have experience with gambling made fewer safe choices. To test Hypothesis 4 we use the Mann-Whitney U test, which gives $p = 0.27$, confirming that there is no statistically significant difference between the two samples. We, therefore, reject the null hypothesis, which states that the group containing participants that indicate experience with gambling make fewer safe choices than the rest of the population.

4.8 Does choice of investment products affect risk aversion?

Table VII: Investment products and safe choices

Product	Mean	n
Commercial Paper	5.33	3
Complex Financial Products	8.00	1
Corporate Bonds	7.00	3
Crowdfunding	6.00	1
Cryptocurrencies	6.00	11
Mutual funds	6.10	10
Retirement funds	5.91	11
Stocks	5.79	14
Treasury Bonds	5.92	13

We asked participants to select which, if any, investment product they own, or have owned in the past from a basket of financial market assets. 38 participants (31%) indicate that they have owned at least one of the products in the basket.

Table VII gives a summary of the investment products in the basket, the number of participants who answered that they own, or have owned them, as well as the mean number of safe choices per product. Apart from the outliers with a small sample size, there seems to be little to no difference between the mean number of safe choices and the different products. Although classifying the level of risk of each of the products in the basket, they have varying levels of risk.

4.9 The effect of household income on risk aversion

We were interested in seeing what effect the different levels of stated household income have on risk aversion in our experiment. Holt & Laury (2002) finds a mildly negative relationship between income and risk aversion. To see if these relationships exist in our experiment, we asked our subjects to indicate their household income by picking from a list containing five intervals. Table VIII in the appendix shows the distribution of subjects across the different intervals. Using a linear regression model, we can investigate the relationship between the number of safe choices and stated household income intervals. The model uses the interval with the most number of observations as its reference (Between 1001 and 2500 per month). It is summarized in Table X in the appendix. There seems to be a negative relationship between the number of safe choices and the income intervals, meaning, as stated household income increases, the mean number of safe choices decreases. However, this effect is only statistically significant ($p < 0.05$) for the income interval between 2510 and 4000 per month. For the other levels, the relationship is unclear as it is negative for incomes that are both lower and higher than the reference level; furthermore, the effect at these levels is not statistically significant. In conclusion, this model does not sufficiently explain the relationship between household income.

5 Conclusion

In this paper, we implement a Holt & Laury (2002) experiment and apply it to a specific cohort of the Portuguese population, namely Generation Z, intending to test whether they exhibit different preferences towards risk than what has been determined in the existing body of literature. Our results are broadly consistent with existing literature on the topic. We find that subjects still reveal risk-averse preferences even faced with lottery tickets instead of monetary prizes as rewards. Additionally, we find that female subjects are slightly more risk averse than male subjects.

The time and recruitment constraints faced when executing the sessions may have led to a biased sample towards students of an economics faculty, which are not representative of the population as a whole. In addition, the recruitment constraints may have led to a self-selection bias in our sample as there were only a few sessions subjects could participate in, with short notice.

Although our implementation had some changes that made it stray from the experiment we aimed to replicate, we were able get a dataset that let us analyse the risk preferences of a sample of the desired population. The change mentioned above relates mainly to the fact that participants were free to withdraw between all rounds, which reduced the quality and increased the effect of noise in the data we had from different treatments in the experiment. This made it difficult for us appropriately analyse the impact of linearly scaling payoffs in the different treatments.

Nevertheless, the technical implementation of this experiment can, with improvements and further considerations, be used in a larger experiment where the research question of this paper is answered by using a representative sample. Further research, should instead of relying on treatments using scale effects, be designed around using a combination of experimental risk preference elicitation methods to get better data and more robust results that could be of use to policymakers.

6 Bibliography

Andersen, S., Harrison, G. W., Lau, M. I., & Rutström, E. E. (2006). Elicitation using multiple price list formats. *Experimental Economics*, *9*(4), 383–405. https://doi.org/10.1007/s10683-006-7055-6

Binswanger, H. P. (1980). Attitudes toward Risk: Experimental Measurement in Rural India. *American Journal of Agricultural Economics*, *62*(3), 395–407. https://doi.org/10.2307/1240194

Camerer, C. F., & Hogarth, R. M. (1999). The Effects of Financial Incentives in Experiments: A Review and Capital-Labor-Production Framework. *Journal of Risk and Uncertainty*, *19*(1), 7–42. https://doi.org/10.1023/A:1007850605129

Charness, G., Garcia, T., Offerman, T., & Villeval, M. C. (2020). Do measures of risk attitude in the laboratory predict behavior under risk in and outside of the laboratory? *Journal of Risk and Uncertainty*, *60*(2), 99–123. https://doi.org/10.1007/s11166-020-09325-6

Charness, G., Gneezy, U., & Imas, A. (2013). Experimental methods: Eliciting risk preferences. *Journal of Economic Behavior & Organization*, *87*, 43–51. https://doi.org/10.1016/j.jebo.2012.12.023

Chen, D. L., Schonger, M., & Wickens, C. (2016). oTree—An open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, *9*, 88–97. https://doi.org/10.1016/j.jbef.2015.12.001

Crosetto, P., & Filippin, A. (2016). A theoretical and experimental appraisal of four risk elicitation methods. *Experimental Economics*, *19*(3), 613–641. https://doi.org/10.1007/s10683-015-9457-9

Dave, C., Eckel, C. C., Johnson, C. A., & Rojas, C. (2010). Eliciting risk preferences: When is simple better? *Journal of Risk and Uncertainty*, *41*(3), 219–243. https://doi.org/10.1007/s11166-010-9103-z

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*(3), 522–550. https://doi.org/10.1111/j.1542-

4774.2011.01015.x

Engel, C., & Kirchkamp, O. (2019). How to deal with inconsistent choices on multiple price lists. *Journal of Economic Behavior & Organization*, *160*, 138–157. https://doi.org/10.1016/j.jebo.2019.02.024

Filippin, A., & Crosetto, P. (2016). A Reconsideration of Gender Differences in Risk Attitudes. *Management Science*, *62*(11), 3138–3160. https://doi.org/10.1287/mnsc.2015.2294

Harrison, G. W., Johnson, E., McInnes, M. M., & Rutström, E. E. (2005). Risk Aversion and Incentive Effects: Comment. *American Economic Review*, *95*(3), 897–901. https://doi.org/10.1257/0002828054201378

Harrison, G. W., Lau, M. I., & Rutstroem, E. E. (2007). Estimating risk attitudes in Denmark: A field experiment. *Scandinavian Journal of Economics*, *109*(2), 341–368. https://doi.org/10.1111/j.1467-9442.2007.00496.x

Holt, C. A., & Laury, S. K. (2005). Risk aversion and incentive effects: New data without order effects. *American Economic Review*, *95*(3), 902–904. https://doi.org/10.1257/0002828054201459

Holt, C. A., & Laury, S. K. (2002). Risk Aversion and Incentive Effects. *American Economic Review*, *92*(5), 1644–1655. https://doi.org/10.1257/000282802762024700

Ihli, H. J., Chiputwa, B., & Musshoff, O. (Eds.). (2016). Do Changing Probabilities or Payoffs in Lottery-Choice Experiments Affect Risk Preference Outcomes? Evidence from Rural Uganda. *Journal of Agricultural and Resource Economics*. https://doi.org/10.22004/ag.econ.235193

Kirchkamp, O., Oechssler, J., & Sofianos, A. (2021). The Binary Lottery Procedure does not induce risk neutrality in the Holt & Laury and Eckel & Grossman tasks. *Journal of Economic Behavior & Organization*, *185*, 348–369. https://doi.org/10.1016/j.jebo.2021.02.026

Statistics Portugal. (2020). Internet and computer usage in households by Nuts II, annual [Dataset]. In *Internet and computer usage in households by Nuts II, annual*. https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_indicadores&userLoadSave=Load&userTableOr

Yu, C. W., Zhang, Y. J., & Zuo, S. X. (2021). Multiple Switching and Data Quality in the Multiple Price List. *The Review of Economics and Statistics*, *103*(1), 136–150. https://doi.org/10.1162/rest_a_00895

A Appendix A

Summary statistics

Table VIII: Summary statistics

Characteristic	**N = 90**
district	
Lisboa e Vale do Tejo/Norte	53 (59%)
Centro	16 (18%)
Autonomous Region of Madeira	2 (2.2%)
Alentejo	8 (8.9%)
Algarve	5 (5.6%)
Other/Foreign	5 (5.6%)
Autonomous Region of the Azores	1 (1.1%)
education	
Undergraduate or Bachelor's degree	38 (42%)
Upper Secondary Education	50 (56%)
Master, MBA or Doctorate degree	2 (2.2%)
work	
Student	79 (88%)
Employed	10 (11%)
Other	1 (1.1%)
risk	
Risk neutral	23 (26%)
Risk loving	28 (31%)
Risk averse	35 (39%)
Very risk averse	1 (1.1%)

Very risk loving	3 (3.3%)
income	
Between 1001 and 2500 per month	36 (40%)
Between 2510 and 4000 per month	26 (29%)
Between 501 and 1000 per month	17 (19%)
More than 4000 per month	10 (11%)
Up to 500 per month	1 (1.1%)
gender	
Female	51 (57%)
Male	39 (43%)
gambling	
Non-gambler	62 (69%)
Gambler	28 (31%)
year	
1997	8 (8.9%)
1998	10 (11%)
1999	11 (12%)
2000	12 (13%)
2001	9 (10%)
2002	11 (12%)
2003	29 (32%)
economics	48 (53%)

[1] n (%)

Withdrawals by treatment

Table IX: Withdrawals sorted by round and treatment

Treatment	Round 1	Round 2	Round 3	Round 4	Total withdrawals by treatment
20x Hypothetical	3	3	1	0	**7**
20x Hypothetical Real	5	1	6	0	**12**
20x Real	2	7	1	0	**10**
50x Hypothetical Real	3	2	6	0	**11**
90x Hypothetical Real	5	5	1	1	**12**
Withdrawals by round	18	18	15	1	**52**

Household income linear regression

Table X: Regression results

	Dependent variable:
	nSafe
incomeBetween 2510 and 4000 per month	-1.019^{**} (0.438)
incomeBetween 501 and 1000 per month	-0.309 (0.501)
incomeMore than 4000 per month	-0.250 (0.608)
incomeUp to 500 per month	2.750 (1.726)
Constant	6.250^{***} (0.284)
Observations	90
R^2	0.095
Adjusted R^2	0.052
Residual Std. Error	1.702 (df = 85)
F Statistic	2.225^{*} (df = 4; 85)
Note:	*p<0.1; **p<0.05; ***p<0.01

B Appendix B

B.1 Instructions

English

Read the following instructions before starting. If you have any doubt, please raise your hand and consult the researcher. Please keep silent during the experience.

In the following screen you will be presented with a series of 10 tasks with two choices. In each choice, we ask you to select **Option A** or **Option B**. In each choice, you can win two prizes, expressed as lottery tickets. For example, if the choice is the following:

> "20% probability of winning 10 lottery tickets and 80% probability of winning 5 lottery tickets."

This means that by selecting this choice, you may win 10 or 5 tickets, where the probability of winning 10 tickets is 20% and 5 with the remaining probability.

Portuguese

Leia as seguintes instruções. Se tiver quaisquer dúvidas levante a mão para chamar o investigador. Deve permanecer em silêncio.

Segue-se uma série de 10 escolhas. Em cada um dos casos, pedimos que escolha **Opção A** ou **Opção B**. Em cada uma das opções, pode ganhar dois prémios, expressos em número de bilhetes de lotaria. Por exemplo, se a opção dá:

> "20% de probabilidade de ganhar 10 bilhetes e 80% de probabilidade de ganhar 5 bilhetes."

Isto significa que ao escolher esta opção, habilita-se a ganhar 10 ou 5 bilhetes de lotaria, ganhando 10 bilhetes com probabilidade 20% e 5 bilhetes com a probabilidade restante.