LUISS Guido Carli
PREMIO TESI D'ECCELLENZA

An Ambiguity Analysis under Heterogeneity with a Bayesian Spin

Diletta Topazio

5
2017-2018

LUISS
Academy

LUISS Guido Carli / Premio tesi d'eccellenza
Working paper n. 5/2017-2018
Publication date: december 2019
An Ambiguity Analysis under Heterogeneity with a Bayesian Spin

ISBN 978-88-6105-493-6

LUISS Academy is an imprint of
LUISS University Press – Pola Srl
Viale Pola 12, 00198 Roma
Tel. 06 85225485
e-mail lup@luiss.it
www.luissuniversitypress.it

An Ambiguity Analysis under Heterogeneity with a Bayesian Spin

Diletta Topazio

Abstract

We often have to deal with uncertainty regarding multiple aspects of the decision problems we face. This uncertainty may concern, for instance, our earnings, the likelihood to receive them in a given moment and in a given amount. The aim of this thesis is to contribute to the growing body of literature around “multi-dimensional uncertainty”, which enlarges the scope of ambiguity outside the frame of uncertainty about probabilities. It does so by analysing, both theoretically and empirically the evidence stemming from a multi-stage experiment in which subjects have to choose between lotteries whereby amounts of monetary prizes are not always known, whereas probabilities are always public knowledge. In the experiment, three different levels of information over some monetary prizes are randomized between subjects. The experimental evidence undergoes structural estimation exercises: these elicit the individuals' degree of risk aversion within the frame of a standard constant relative risk aversion (CRRA) utility function. Furthermore, we investigate whether a change of information, such as the one we reproduce through the different treatments conditions, translates into a change in behavior and, in turn, whether and how much this change translates into a significant change in their measured (CRRA) attitude toward risk. As to the behavioral content of the structural model for the uncertain payoffs, we propose two alternative specifications, labelled “naïve” and “sophisticated”. The empirical evidence shows a moderate but significant degree of *love for ambiguity*, since less information given to subjects results in a lower estimate of their risk aversion, and, as a consequence, in a stronger attraction toward risk and uncertainty. A mixture model is implemented to identify the probability of individuals mirroring one behavioral model or the other, or, saying it differently, the percentage of observations compatible with either model. We conclude that our subjects have a strong tendency to behave as naïve.

Keywords: heterogeneity; risk aversion; ambiguity

1. Introduction

Uncertainty regarding multiple aspects of the decision problems is an issue that often arises. This uncertainty may concern the amount of monetary earnings, the likelihood of these earnings, the actual date at which these earning are received, etc. In this respect, the way in which individuals behave in uncertain situations may as well varies for different dimensions of uncertainty. However, the theoretical and empirical economic discussion on these issues has been mostly focused on a specific kind of uncertainty, the *uncertainty about probabilities*.

The aim of this dissertation is to contribute to the scarce but increasing body of research which deals with "multi-dimensional uncertainty", that enlarges the scope of ambiguity outside the frame of uncertainty about probabilities. We contribute on this by analyzing (both theoretically and empirically) existing evidence from a multi-stage experiment in which subjects have to choose between lotteries where probabilities were publicly known at all times but some of the monetary prizes were not.

It should be noted at this point that, from a pure bayesian perspective, all the different dimensions of ambiguity can be reduced to a single one by appropriately defining the "states of the world" as multidimensional objects defined over all uncertain dimensions. Within this augmented frame, a "bayesian" decision maker would simply form some subjective prior beliefs over this augmented set of states of the world and maximize an objective function, that represents her preferences based on them. If we accept this bayesian interpretation, every decision problem under multi-dimensional ambiguity can be appropriately reduced to a standard problem of uncertainty over probabilities. This can only stand if we show that people are able to build such complex and multi-dimensional spaces and behave accordingly. Conversely, if this was not be the case, it would matter which objects of uncertainty are domains and whether and how these domains might be correlated.

This thesis reports evidence form a multi-stage experiment conducted at the "Laboratory of Theoretical and Experimental Economics" of the University of Alicante by Albarrán et al [1]. In the experiment, three different levels of information over some monetary prizes are randomized between subjects. Specifically, in the full information treatment, TR2, subjects observe all the prizes of the lotteries they are asked to select; in the partial information treatment, TR1, they are not informed about their actual values, but they know that they are i.i.d. draws from a uniform distribution; finally in the no information treatment, TR0, they are just informed of the prize rankings.

The experiment develops along two ordered balance phases built upon two classic risk-elicitation protocols, the Holt and Laury [2] and the Hey and Orme [3], respectively. Hey and Orme [3] experiment is built around binary choices between lotteries over 4 fixed monetary prizes, such as $\{0, 1/3, 2/3, 1\}$. In the treatments with ambiguity, TR0 and TR1, the intermediate payoffs, Y and X, are communicated to the subjects, being between 0 and 1, with $Y<X$. In phase 1, instead, subjects elicit, by the way of a Multiple Prize List (MPL), the certainty equivalent of the same lotteries used in phase 2.

The experimental evidence is read by the way of some structural estimation exercises in which the individuals' degree of risk aversion is elicited within the frame of a standard constant relative risk aversion (CRRA) utility function. Furthermore, it is analyzed whether a change of information, such as the one reproduced through the different treatments conditions, correlates to a change in behavior and, in turn, whether and, how this change transforms into a significant change in their measured (CRRA) attitude toward risk.

The uncertain payoffs Y and X are identified as the first and the second order statistics from a uniform distribution in [0, 1], where the order statistics of a random sample $\{X_1, X_2\}$ are defined as the sample values placed in ascending order.

While this is totally correct for TR1 subjects -given that they know the characteristics of the random generation process that yields the uncertain payoffs- the same statistical model was imposed for subjects in TR0, considering that they already had this information. This is purely an identification assumption, as there is no possibility to test whether this is truth for the expectations in TR0 about the X and Y distributions, or whether subjects in TR0 consider another distribution. On the other hand, it is highly probable that TR0 subjects will heuristically and automatically come up with such a distribution of the payoffs, as it occurs in Laplace's well known "principle of insufficient reason". In any case, the important here is that -based on this assumption- our structural model is able to estimate treatment effects, to such an extent that we are able test a null hypothesis in which CRRA in both TR0 and TR1 is the same. Since subjects are randomized within treatments, a significant change in the CRRA coefficient between TR0 and TR1 has to be interpreted as a genuine treatment effect due to a change in information.

Regarding the behavioral content of the structural model for the uncertain payoffs, two alternative specifications were considered, labelled as "naïve" and "sophisticated", respectively. A decision maker figures out a point estimation of the unknown payoffs *X* and *Y*, starting from the information that these are draws of a uniform distribution in [0, 1]. This means that E[X] and the E[Y] are computed and then plugged inside the CRRA utility function to be maximized. On the other hand, a sophisticated decision maker will proceed with a true bayesian updating. In particular, she formulates a prior distribution over the X and the Y, and then calculates the expected utility from these densities.

We shall now here summarize our main findings. Our empirical evidence shows a certain degree of *love for ambiguity*, given that the less the information given to subjects, the lower their estimated risk aversion, and, consequently, the bigger their attraction toward risk and uncertainty. Moreover, the risk aversion coefficient estimated for TR0 is significantly lower from that estimated in TR2, although no statistically significant difference was found between estimated CRRA coefficients in TR0 and TR1. These findings are in a way in contradiction to the common wisdom of the literature, although they are consistent with other experimental literature that applies similar elicitation techniques as Andersen et al [4].

When comparing the two behavioral models, that is the "bayesian" against the "naïve", the estimated likelihood of the naïve approach is higher than the one of the bayesian. This suggests that the naïve approach closely approximates subjects' decision rules, given the data.

In this regard, a mixture model is implemented to identify the probability of individuals using each model, which corresponds to the percentage of observations compatible to each model. This mixture model aims to achieve a statistical reconciliation of these two dominant theories of choices under risk. It avoids any extreme declaration of "winners" and "losers", providing a more balanced metric to decide which theory performs better in a given domain given the experimental data. Due to the fact that the likelihoods of our models are very close, this

probability was estimated numerically, using a grid loop.
Specifically, a probability π_{BAY}, i.e. the probability of the subjects acting as bayesian in each of their decisions was estimated. Subsequently we let this π_{BAY} moving inside a grid (0,1), to finally choose the value that maximizes the likelihood function.
This numerical computation demonstrates that the subjects have a strong tendency to behave as naïve, given the estimation result, which was $\pi_{BAY} = 0.2$.
The structure of this thesis is arranged as follows. In Chapter 2 it is shown how the individuals' heterogeneity is treated under structural modelling, providing some examples applied to the data. In Chapter 3 our structural estimations are reported as a function of the two alternative behavioral specifications, the naïve and the bayesian, along with the estimation of our mixture model. Finally, Chapter 5 summarizes our results and highlights possible future developments and more complex experimental investigations.

2. Structural Modelling Under Heterogeneity

2.1 Maximum Likelihood Estimation

Customized likelihood functions corresponding to specific models of decision making under risk and uncertainty are more and more popular among economists dealing with a wide range of fields, as suggested by Harrison [29]. This demand for customization is due to the numerous parametric functional forms experimental economists use to account for behaviour under risk and uncertainty. These functions also permit to represents ”handwritten” models, used to explain decision rules, which may be different from the traditional ones. In behavioural econometrics it's becoming even more common to see user-written maximum likelihood estimations rather than pre-packaged model specifications.
Specifically, what a maximum likelihood estimation does is, conditional on the structural model under scrutiny, to select the value of the estimator which maximizes the probability of observing the collected data, i.e. the probability density function. It is given by a model such as $P(y \mid \Theta)$, where Θ represents the set of unknown parameters we want to estimate and Y the vector containing the observed decisions. The maximum likelihood estimator, Θ^*, maximizes the likelihood function $P(y \mid \Theta)$ with respect to Θ; this means that we maximize the probability of observing the data we actually observed as function of the parameters of the model.
With a big sample size the likelihood function, being the product of the probability density functions of all the subjects' outcomes, is close to 0. For this reason, locating its maximum may be difficult. The logarithmic function is usually employed to solve this problem in order to stretch the function vertically and making it easier to locate its maximum. Furthermore, the logarithmic transformation is strictly monotone, preserving the same local maxima.
In our specific model, so under the Expected Utility Model, the probabilities of each

outcome k, namely p_k, are the ones induced by the experimenter[1]. This means that the expected utility is calculated as the sum of utilities of monetary prizes, each of which is multiplied by the corresponding probability.

Being i the subject, $k \in \{0, 1\}$ the index of the lottery equal to 0 for the right one and 1 for the left one, and $h \in \{1, 2, 3, 4\}$ the index of the prizes:

$$\boldsymbol{U_i(L_k) = \sum_{h=1}^{H} u(\chi_i)\, p_h^h} \tag{2.1}$$

where $\mathbf{u}_i : \mathrm{R} \rightarrow \mathrm{R}$ $\quad L_k = \{X; p_k\}$ $\quad \boldsymbol{X \in \{0, y, x, 1\}}$

In this chapter we use the mean-variance utility function (2.2), given its simplicity and intuitiveness in calculations. The mean-variance (MV) utility function applied to a lottery i is as follow:

$$\mathrm{U_i(L_k)} = E[\mathrm{L_k}] - \beta_i\, Var(\mathrm{L_k}) \tag{2.2}$$

The utility of each individual is a function of the expected value of the lottery μ_k, and of its variance, $Var(\mathrm{L_k})$.

It can be shown that MV utility is equivalent to a VNM utility function (2.1) with a quadratic utility function $\mathrm{u(x) = x - \beta x^2}$, where β is the only unknown parameter to be estimated. It represents the level of risk aversion. Indeed, the variance $Var(\mathrm{L_k})$ is used as a proxy of the risk of the lottery. An individual is considered to be risk-averse if β has a positive value, namely if, ceteris paribus, a higher value of the lottery variance decreases his utility. By the same token, a negative β value is associated with a risk-seeking decision maker, hence β =0 indicates risk neutrality.

In this section we shall estimate β using data from the full information treatment, TR2, where there is no ambiguity and subjects know the true value of both X and Y.

In Phase 1, subjects make 16 choices per period. According to the usual Holt and Laury framework, the threshold value of β, such that the decision maker switches to the certain amount rather then the risky lottery, is extracted. According to Moffatt [30], only this piece of information should be extracted from each individual facing the Holt and Laury lottery. Indeed, he states the list of people choices cannot be analyzed as an independent sequence. However, this does not seem to be the case for our data, as we can see from Figure 2.1 from Rodriguez and Ponti [31], where subjects choices are shown as a function of the periods (i.e., the individual choice between the lottery and a specific monetary prize ranging from 0 to 15).

[1] All maximum likelihood routines have been programmed and run with STATA v. 14, by STATA Corporation.

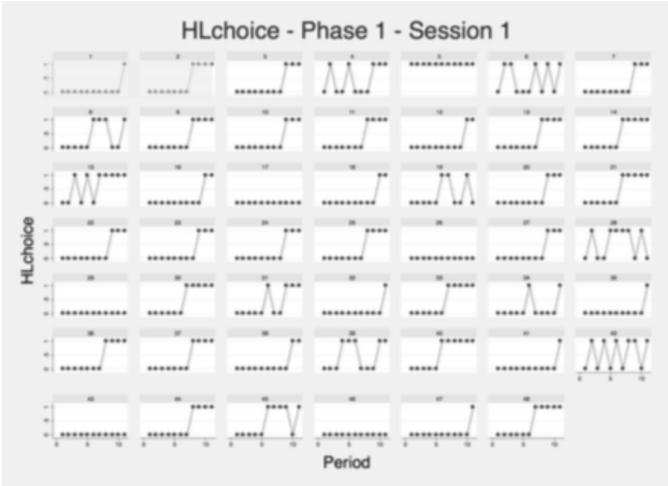

Figure 2. 1: Phase 1 Switching Point Graph

Here the first two highlighted subjects, subject 1 and 2, are a perfect example of a rational behavior, as indicated by the presence of a single switching point. Looking at subject 4 or 6 instead, a clearly irrational behavioral pattern emerges, with the presence of several switching points. Furthermore, these two cases are not unique. This demonstrates that we may not rely on a single switching point information per subject. Indeed, the emergent absence of a clear path individuals follow throughout the whole Phase suggests that we may also treat people's decisions as independent.

For Phase 2 we directly take the 25 decisions subjects make between the two lotteries.

For the lotteries of Phase 1 and Phase 2, the expected value is computed as:

$$E[L_k] = \sum_{h=1}^{H} p_k^h \times X_k,$$

where π_h is the prize the subjects will receive with probability p_k.

The variance, as the expectation of the squared deviation of the prize random variable from its mean, is computed as:

$$\sigma_{L_i}^2 = \sum_{h=1}^{H} p_k^h \times (\chi_h - E[L_k])^2.$$

Then, using these values, for every left and right lottery we compute their utility function and the difference of these utility values:

$$U_i(L_k) = E[L_k] - \beta_i \, \sigma_{L_k}^2, \;\; k = 0{,}1.$$

This difference will prescribe the optimal behavior, namely the lottery to be chosen according to the utility maximization principle.

In Phase 1, the left lottery L_0 is actually a certain prize. This means that for all the left lotteries we have $\sigma_0^2 = 0$ and $E[L_0] =$ the actual number displayed. Namely, the β never appears in the left lotteries of Phase 1.
Subsequently, we proceed with the calculation of the difference of the expected utility, as follows:

$$\Delta U = U(L_1) - U(L_0) = \left(E[L_1] - \beta\sigma_{L_1}^2\right) - \left(E[L_0] - \beta\sigma_{L_0}^2\right) = [\dots]$$

This latent index, based on latent preferences, is then brought back to the observed choices using a standard cumulative normal distribution function $\Phi(\Delta E[U])$.
This probit function has a domain in $[-\infty, +\infty]$ and has a codomain in [0,1] as shown in Figure 2.2.

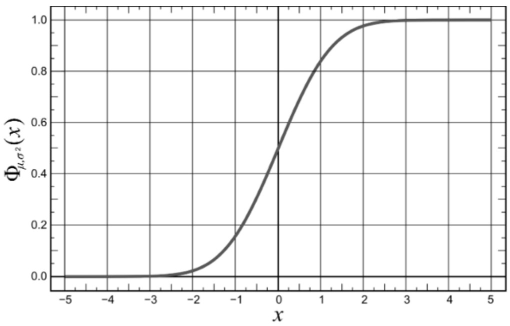

Figure 2. 2: Normal c.d.f Function

2.2 Dealing with Heterogeneity
Here we introduce the distinction between the various and competing approaches to stochastic modelling, in which the choices made by individuals express their heterogeneity.
The Random Preference Model (Loomes and Sugden [32]) explains heterogenity in decisions as heterogeneity in the structural component β, in our case). This is an example of an heterogeneous agent approach, which attributes a variation in behavior of the population to variation in the parameter representing preferences.
The Fechner Model (Fechner [33]), in which the stochastic component in the decision making process is done applying an additive idiosyncratic error $\epsilon \sim N(0, \sigma_\epsilon^2)$.
The Tremble Approach (Loomes et al. [34]) assumes that there is a small positive probability that the individual, at any point in time, loses concentration and adopt any possible behavior randomly with probability ω.
The Random Effect Model, whereby the unobserved heterogeneity is expected to be explained by a random effect parameter, η_i, which captures the between subject differences, and its variance, σ_η^2, is a measure of subject heterogeneity.

Random Preference Model
In this model, the structural parameter $\beta \sim N\left(\mu_\beta, \sigma_\beta^2\right)$ accounts for all the heterogeneity in the model.
For the maximum likelihood routine to work, we need to create some local variables, which

are temporarily used only inside the program.
The parameters we want to estimate are μ_β and σ_β^2, which describe the distribution of $\beta \sim N(\mu_\beta, \sigma_\beta^2)$. This means that we maximize the log likelihood function with respect to μ_β and σ_β^2, where:
- μ_β is estimated through the choices made by subjects, the experimental data represents then the basis to construct likelihood functions, such that the selected optimal value of β is directly dependent on the subjects' binary decisions between the two lotteries.
- σ_β^2 is imposed to be strictly positive, given that it is the standard deviation of β, using the strictly monotone exponential transformation, and is estimated as a constant.
Equating the expected utilities and solving for β we get:

$$\beta = \frac{\mathrm{E}[\mathrm{L}_1] - \mathrm{E}[\mathrm{L}_0]}{\sigma_{\mathrm{L}_1}^2 - \sigma_{\mathrm{L}_0}^2}$$

This general formula holds for both Phase 1 and Phase 2. Furthermore, for Phase 1, it can be simplified, given the left lottery is always a fixed price, γ, and $\mathrm{E}[\mathrm{L}_0] = \gamma$ and $\sigma_{\mathrm{L}_0}^2 = 0$.
Specifically, $\gamma_\mathrm{i} = \frac{\delta_\mathrm{i}-1}{15}$, with δ_i being the prize of decision *i*, is scaled back by 1 position and then normalized in [0,1] dividing by 15.

$$\beta_{\mathrm{Phase1}} = \frac{E[L_1] - \gamma}{\sigma_{\mathrm{L}_1}^2}$$

Therefore, we are left with these two values for β:

$$\beta_{\mathrm{Phase1}} = \frac{E[L_1] - \gamma}{\sigma_{\mathrm{L}_1}^2}$$

and

$$\beta_{\mathrm{Phase2}} = \frac{E[L_1] - E[L_0]}{\sigma_{\mathrm{L}_1}^2 - \sigma_{\mathrm{L}_0}^2}$$

The probability that the right lottery is chosen, conditional on the average risk aversion coefficient, namely the mean of β, μ_β, is equal to the probability that ΔU is positive:

$$\mathrm{P}(\mathrm{k} = 1 | \mu_\beta) = \mathrm{P}(\Delta\mathrm{U} > 0 \mid \mu_\beta) = \mathrm{P}(\mathrm{U}(\mathrm{L}_1) > \mathrm{U}(\mathrm{L}_0) \mid \mu_\beta)$$
$$= \mathrm{P}(\Delta\mathrm{E}[\mathrm{L}_\mathrm{k}] - \beta\Delta\mathrm{E}[\sigma_\mathrm{k}] > 0 \mid \mu_\beta) = \mathrm{P}\left(\beta < \frac{\Delta\mathrm{E}[\mathrm{L}_\mathrm{k}]}{\Delta\mathrm{E}[\sigma_\mathrm{k}]} \middle| \mu_\beta\right).$$

Calling $\beta^* = \frac{\beta-\mu_\beta}{\sigma_\beta}$, we use the normal transformation and we get:

$$P_{i,t}\left(\beta < \frac{\Delta E[L_{k,t}]}{\Delta E[\sigma_{k,t}]} \,\middle|\, \mu_\beta\right) = P\left(\frac{\beta - \mu_\beta}{\sigma_\beta} < \frac{\frac{\Delta E[L_{k,t}]}{\Delta E[\sigma_{k,t}]} - \mu_\beta}{\sigma_\beta} \,\middle|\, \mu_\beta\right) = P(\beta < \beta^* \mid \mu_\beta)$$

$$= \Phi(\beta^*_{i,t})$$

where Φ is the standard normal c.d.f.

In conclusion, the likelihood function of observing the right lottery chosen, given μ_β, is:

$$\mathcal{L} = \sum_{i,t} \ln\left(P_{i,t}\left(\beta < \frac{\Delta E[L_{k,t}]}{\Delta E[\sigma_{k,t}]} \,\middle|\, \mu_\beta\right)\right)$$

An important property of the c.d.f. of both the probit and the logit model is symmetry.

By this, $\Phi(-\beta^*) = 1 - \Phi(\beta^*)$.

Following the same way of reasoning, the symmetric probability that the right lottery is chosen, is just:

$$P(k = 0 \mid \mu_\beta) = P(\Delta U < 0 \mid \mu_\beta) = P(U(L_1) < U(L_0) \mid \mu_\beta)$$

$$= P(\Delta E[L_k] - \beta\Delta E[\sigma_k] < 0 \mid \mu_\beta)][= P\left(\beta > \frac{\Delta E[L_k]}{\Delta E[\sigma_k]} \,\middle|\, \mu_\beta\right)$$

Calling $\beta^* = \frac{\beta - \mu_\beta}{\sigma_\beta}$, we use the normal transformation and we get:

$$P_{i,t}\left(\frac{\beta - \mu_\beta}{\sigma_\beta} > \frac{\frac{\Delta E[L_{k,t}]}{\Delta E[\sigma_{k,t}]} - \mu_\beta}{\sigma_\beta} \,\middle|\, \mu_\beta\right) = P_{i,t}(\beta > \beta^* \mid \mu_\beta) = \Phi(-\beta^*)$$

In conclusion, the likelihood function of observing the left lottery chosen, given μ_β is:

$$\mathcal{L} = \sum_{i,t} ln\left(1 - P_{i,t}\left(\beta < \frac{\Delta E[L_{k,t}]}{\Delta E[\sigma_{k,t}]} \,\middle|\, \mu_\beta\right)\right)$$

Fechner Model

As hinted above, the Fechner Model introduces a stochastic component to take into account subjects' heterogeneity in the decision making process. This is included by adding an idiosyncratic error to ΔU, that is $\epsilon \sim N(0, \sigma_\epsilon^2)$. Then, rather than to parameters, attention is here posed to payoffs differences.

As soon as this error terms appears, the behavior is no longer deterministic and it is described in term of probabilities as follows:

$$P(k=1) = P(U(L_1) + \xi_1 > U(L_0) + \xi_0) = P(\Delta U + \epsilon > \mathbb{0}) = P(\epsilon > -\Delta U)$$
$$= P\left(\frac{\epsilon}{\sigma_\epsilon} > \frac{-\Delta U}{\sigma_\epsilon}\right) = \Phi\left(\frac{\Delta U}{\sigma_\epsilon}\right)$$

where ϵ is the difference between the two lotteries' errors in valuation ξ_1 and ξ_0, Φ is the standard normal c.d.f. and σ_ϵ represents the noisiness of the choice. This means $\sigma_\epsilon = 0$ fully explains a deterministic choice, while if $\sigma_\epsilon \to \infty$, the choice is entirely driven by noise, namely both right and left lotteries are chosen with 0.5 probability.
The parameter we want to estimate is σ_ϵ, namely we maximize the log likelihood function with respect to it. It is imposed to be strictly positive, being a standard deviation, trough the usual strictly monotone exponential transformation.
Again, the $U(L_i)$and the ΔU are used as temporary variables.
Table 2.1 reports an application of the Fechner model using the Mean Variance utility function. Only TR2 data are used here, so β estimated is a proxy of individuals' aversion to risk (as represented by the variance).

VARIABLES	
Beta	0.458***
	(0.0827)
Sigma	-2.033***
	(0.0644)
Obs.	61,339

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Tabel 2.1: Fechner Model with Mean Variance Utility Function

Tremble Parameter

According to the Random Preference model just described above, if the lottery L_i first-order stochastically dominates lottery $\$L_k$, the first one will always be chosen, no matter the risk attitudes of subjects. Indeed, any observed choice of a dominated lottery cannot be explained by the RP model.
For this reason, the tremble parameter ω is introduced, and it represents the probability a subject loses concentration at any task and randomly chooses, with equal probability, between the two alternative lotteries. This does not necessarily imply he makes the incorrect choice, as

under such a condition the correct and incorrect choices are equally likely.
This is not a model *per se*, rather it's an extension to be applied either to the Random Preference model or to the Fechner model.

In the RP model the $P(k=1)(1-\omega)\Phi\left(\frac{\beta^{*}-\mu_{\beta}}{\sigma_{\beta}}\right)+\frac{\omega}{2}$.
In Fechner, $P(k=1)=(1-\omega)\Phi\left(\frac{\Delta U}{\psi}\right)+\frac{\omega}{2}$.

This means that with (1 - ω) probability the correct choice prescribed by the model is done, and with ω probability the random choice equal to $\frac{1}{2}$ is made.

The Random Effect Model
In this model, the unobserved heterogeneity is explained by a random effect parameter, η_i, which is perpendicular to the other covariates. Specifically, η_i captures the between-subject differences, and its variance, σ_η^2, is a measure of subject heterogeneity. The probability of observing the right lottery chosen by subject *i* in period τ is now:

$$P(k=1)=f(\chi_{i,\tau}\,\beta)+\eta_i+\epsilon_{i\tau}$$

where $\eta_i \sim (0,\sigma_\eta^2)$ is the individual error, or heterogeneity, and $\epsilon_{i\tau} \sim N(0,1)$ is the idiosyncratic error. While $\epsilon_{i\tau}$ varies across subjects and periods, η_i has a unique value for every individual.
In our model, this probability can be explained as:

$$P(k=1)=\alpha E\left[\Delta\mu_{i,\tau}\right]+\beta E\left[\Delta\sigma_{i,\tau}^{2}\right]+\eta_i+\epsilon_i,$$

where

$$E\left[\Delta\mu_{i,\tau}\right]=E[L_1]-E[L_0]$$

and

$$E\left[\Delta\sigma_{i,\tau}^{2}\right]=\sigma_{L_1}^2-\sigma_{L_0}^2$$

If the α = 1 constraint is imposed, we obtain the usual mean-variance utility function, where a negative value for β is expected.
Only the data from TR2 individual is used, since we want to extract the risk aversion coefficient, while leaving aside any form of ambiguity which might arise from the missing payoffs' information of the TR1 and TR0.
We implement both probit and logit regressions, with the xtprobit and xtlogit functions, namely declaring our data structure to be a longitudinal panel. The Fechner approach is implemented by default when using these functions in Stata. \\The results are shown in table 2.2.
The results of the probit and the logit are closely comparable. As expected, β is negative, and

statistically significant at the 99% confidence level. Namely, the probability of choosing the right lottery is a negative function of the difference between the right and the left lottery variance.
Also the standard deviation of the random effect coefficient is statistically significant at the 99% confidence level, and it gives us information about the subjects heterogeneity. Specifically, the random effect coefficient $\eta_i \sim (0{,}0.238)$ according to the probit model, and $\eta_i \sim (0{,}0.363)$ according to the logit.

VARIABLES	Probit	Logit
Delta_VAR	-0.464***	-0.355***
	(0.0525)	(0.0811)
Random Effect SD	0.238***	0.363***
	(0.011)	(0.0169)
Obs.	195,459	195,459
Sbj.	279	279

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table 2.2: Random Effect Probit and Logit Models

3. Dealing with Ambiguity

3.1 Dealing with Heterogeneity

In conditions of uncertain outcomes, the Savage approach [7] has been traditionally used. In particular, individuals have been assumed to behave according to a unique subjective prior belief over all states of the world, and, given this, they would maximize their expected utility. This decision process clearly neglects the existence of any form of ambiguity, and it prescribes the way decision makers should deal with uncertain situations.
However, Ellsberg [8] claims that most individuals treat ambiguity differently than objective risk. In specific, he argues that people exhibit a significant degree of *ambiguity aversion*, placing a premium on outcomes for which probabilities are known. This general stylized fact has been replicated broadly and has important implications for the economics of optimal contracting, investment choices, and mechanism design.
One possible way to structurally identify ambiguity aversion is to assume that the latter influences people's degree of risk aversion (more precisely, the curvature of the utility function), an approach followed, among others, by Klibanoff et al. [11] and Andersen et al. [4]. As described in Chapter 3, in the experiment of Albarrán et al [1], prizes in the lotteries are distributed according to the rule $0<y<x<15$. In what follows, this prize domain is normalized, for the sake of simplicity, to lay within the unit interval [0, 1], where \$0 is 0 and \$15 is 1.
The treatment conditions -randomized between subjects- regard the amount of information given to them about X and Y. Furthermore, while in the full information treatment, TR2, people face a normal risky situation and there is no ambiguity influencing their decision, this is not

the case for the partial information and no information treatments TR1 and TR0, respectively. As we shall see, some ambiguity preference appears from subjects' choices which is higher the less information is received.

3.2 Econometric Strategy

In what follows we shall layout the identification assumptions underlying our structural estimations. Specifically, we need to define our identification strategy with respect to *i)* subjects' risk attitudes and how the uncertain payoffs, X and Y, enter in subjects' calculations together with *ii)* the behavioral model underlying subjects' optimization program. Regarding the former, as it will be explained in Section 3.2.1 we shall impose that subjects maximize a VNM CRRA utility function in all treatments and that, consistently with the TR1 experimental instructions, Y and X are calculated as first and second order statistics of a uniform distribution defined over the unit interval. Regarding the latter, that is explained in Section 3.2.2, we shall consider two alternative behavioral models, defined as *naïve* and *sophisticated*. In the former, subjects are assumed to estimate first the uncertain payoffs and then use these expected payoffs in the expected maximization program; in the latter -consistently with a genuine bayesian approach- the order of integration is reversed.

3.2.1 Uncertain Payoffs and Risk Aversion

We read the experimental evidence by the way of some structural estimation exercises in which we elicit the individuals' degree of risk aversion within the frame of a standard constant relative risk aversion (CRRA) utility function, which generally performs better in more complex structural estimations.

The utility function is given below:

$$\mathrm{U}(\chi) = \begin{cases} \dfrac{\chi^{1-\rho}}{1-\rho} & if\ \rho \neq 1 \\ \ln \chi & if\ \rho = 1 \end{cases}$$

where ρ is the (CRRA) coefficient which does not depend on χ, as formalized by Pratt [35]. As for its economic interpretation, $\rho > 0$ represents risk aversion, $\rho = 0$ risk neutrality and $\rho < 0$ risk loving.

In Figure 3.1[2], examples of $u(\chi)$ are presented for different values of ρ: concave in case of risk aversion (left) and convex in case of risk loving (right).

[2] Machine [36]

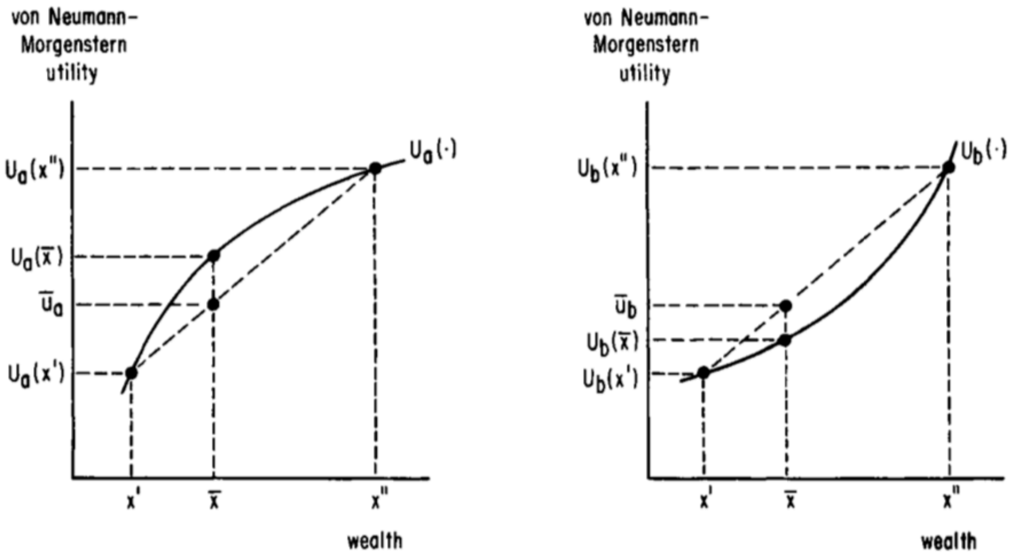

Figure 3.1: Risk Aversion Coefficient for both Risk Aversion and Risk Loving

In Section 3.3.1 we check whether the change of information implemented by our treatments conditions generates a change in behavior and, in turn, (whether and) how this change is converted into a significant change in the measured (CRRA) attitude toward risk.
The uncertain payoffs Y and X are identified as the first and the second order statistics from a uniform distribution in [0, 1], where the order statistics of a random sample $\chi_1, \ldots, \chi_n$ are defined as the sample values placed in ascending order.
Specifically, let $f_k(n, z)$ denote the k_{th} order statistics of *n* draws, where *n = 2* in our case, of a density function $f(\cdot)$.
Let $X \sim f_2(2, z)$, $Y \sim f_1(2, z)$
where

$$f_2(2, z) = 2\, z\, f(z)\, F(z), \qquad f_1(2, z) = 2\, z\, f(z) 1 - F(z)$$

come from the general formula for the k_{th} order statistics of *n* draws

$$n \binom{n-1}{k-1} f(z) \left(F(z)\right)^{n-1} \left(1 - F(z)\right)^{n-k}$$

being *z* a random draw from a uniform distribution and being

$f(z) = 1$ the p.d.f. of *z*, $F(z) = z$ the c.d.f of *z*.

While this is certainly true for TR1 subjects -since they know the characteristics of the random generation process that yields the uncertain payoffs- we impose the same statistical model for subjects in TR0, assuming they had this information. As we said, this is purely an identification assumption, as there is no possibility to test whether this is the true for expectations in TR0 about X and Y distribution, or whether subjects in TR0 consider a different type of distribution. On the other hand, it is highly probable that TR0 subjects will heuristically and automatically assume such a distribution of the payoffs, as it occurs in Laplace's well known "principle of insufficient reason'". In any case, what is important here is that -thanks to this assumption- our structural model is able to estimate treatment effects, to such an extent that we are able test a null hypothesis in which CRRA in both TR0 and TR1 is the same. Since subjects have been randomized within treatments, a significant change in the CRRA coefficient between TR0 and TR1 has to be interpreted as a genuine treatment effect due to a change in information.

In the maximum likelihood function routine, ρ is analyzed through the individual choices subjects make, which are expressed in function of the treatments, to identify how a different level of information influences people's risk attitude.

Phase 1 observations are treated as a series of individual and independent choices between a certain outcome and a risky lottery, whose expected value is computed and compared to the sure prize.

Instead, phase 2 data are used as a sequence of binary choices between lotteries. TR2 players know the true X and Y, so their ρ derived from a situation with no ambiguity. On the other hand, TR0 and TR1 players compute the lotteries expected values and variances, as function of the X and Y they figure out, and then the U_i and the ΔU.

A logit function is used to solve the usual binary choice model, explaining the $P(k = 1) = P(\Delta U > \mathbb{0})$ which is:

$$P(k = 1) = \frac{e^{\Delta U}}{1+e^{\Delta U}} \qquad \text{if } L_1 \text{ is chosen}$$

$$P(k = 1) = \frac{e^{-\Delta U}}{1+e^{-\Delta U}} \qquad \text{if } L_0 \text{ is chosen.}$$

The Fechner model is used, where people heterogeneity is expressed as function of a random error in the CRRA utility computation, i.e. $\epsilon \sim N(0, \sigma^2)$. In the whole of estimates we cluster all the observations made by the decisions of the same individual.

3.2.2 Identification of the Behavioral Model

Regarding the behavioral content of the structural model for the uncertain payoffs, we consider two alternative specifications, labelled as "naïve" and "sophisticated", respectively, naïve. This means that the E[X] and the E[Y] are computed first and then plugged into the CRRA expected utility function to be maximized.

Specifically:

$$E[\,X\,] = \int_0^1 f_2(2,\ z)\, dz = \frac{2}{3}$$

$$E[\,Y\,] = \int_0^1 f_1(2,\ z)\, dz = \frac{1}{3}$$

where f_k is the *k*-th order statistics of a uniform distribution in [0, 1].
Finally, the utility of a lottery *k* is:

$$U(L_k) = u(0)\, p_0^k + u(E[\, Y\,])\, p_y^k + u(E[\, X\,])\, p_x^k + u(1)\, p_1^k$$

A "sophisticated" decision maker, instead, will proceed based on a true bayesian updating, forming a prior distribution over the X and the Y, and then calculate the expected utility from these densities. Specifically:

$$U(X) = \int_0^1 u(z) f_2(2,\, z)\, dz$$

$$U(Y) = \int_0^1 u(z) f_1(2,\, z)\, dz$$

where $f_1(\cdot)$ and $f_2(\cdot)$ are the first and second order statistics of a uniform distribution in [0, 1].
Finally, the utility of L_k, $U(L_k)$ equals to:

$$U(L_k) = U(0)\, p_0^k + U(Y)\, p_y^k + U(X)\, p_x^k + U(1)\, p_1^k$$

In conclusion, the two models differ due to the order of integration.

3.3 Results

The "atom" of our analysis is the decision made by subjects and our research question is how their ρ varies as function of the amount of information they receive, depending on their treatments, and how this process differs in the two distinct approaches, the naïve and the bayesian one. We also query whether one model is more used than the other.

3.3.1 Treatment Effects

Table 5.1 reports the result of the structural estimation of the ρ as function of the different treatments, for both the two approaches.
Our empirical evidence shows a certain degree of *love for ambiguity*, as the less information given to the subjects, the lower their risk aversion, and, consequently, the bigger their attraction toward risk and uncertainty. Moreover, the risk aversion coefficient estimated for TR0 is significantly lower than that estimated in TR2, although there is no statistically significant difference between estimated CRRA coefficients in TR0 and TR1. These findings are -somewhat- in contradiction with the common wisdom of the literature, although they are consistent with other experimental literature that applies similar elicitation techniques as ours, such as Andersen et al [4].
When comparing our two behavioral models, as shown in Table 3.1, the estimated likelihood of the naïve approach is higher than that of the bayesian. This suggests that, based on our data, the naïve approach approximates better subjects' decision rules.
Afterwards, we would like to identify the percentage of the subjects using each of the two models, i.e. the probability of them behaving either in a naïve or a bayesian way.

VARIABLES	Naive	Bayesian
TR_0	-0.0360***	-0.0452***
	(0.00953)	(0.0102)
TR_1	-0.0187*	-0.0254**
	(0.00973)	(0.0102)
Constant	0.774***	0.774***
	(0.00557)	(0.00557)
TR_1 – TR_0	-0.01723	-0.0198
	(.01111)	(0.0120)
Likelihood	-92261.238	-92604.987
Obs.	195,459	195,459

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3.1: Risk Aversion Coefficient for both the Naive and Bayesian approaches

3.3.2 Naïve or Sophisticated?

Up to now, we identified two different approaches individuals may follow to make their choices. The following step is to implement a mixture model to identify the probability of each observation being compatible with either model.

We use a binary mixture model, since a finite number of types, the naïve and the bayesian, are assumed[3].

The main advantage of this approach that the assumption of different subjects operating according to a single model is avoided. The behavior of a typical subject is often traced back to the average behavior, but it is quite possible this is not an accurate representation of every subject under study.

A possible answer to this issue could be the Average Treatment Effect, ATE, where a specific treatment effect is recognized to each individual. All subjects specific treatment effects are then assumed to vary randomly around an average, the ATE, i.e. the parameter being estimated.

If the distribution is bell-shaped and symmetric, the ATE will provide a sensible measure of the effect of the treatment. In other words, the ATE measure is relevant when the treatment has universal applicability so that it is reasonable to consider the hypothetical gain from treatment to a randomly selected member of the population.

However, this is not always the case, and this ATE can end up being far away from the actual treatment effect of any single subject.

The approach adopted by a finite mixture model is presented below. A total number of types in the population is decided, and a specific behavioral model is assigned to each of them. The parameters of these various models are estimated altogether, along with the mixing proportions.

[3] In case of an "infinite" mixture model, a continuous variation in some parameters indexing individual type is assumed, as happens for random coefficient models or random effect models

In particular, we generate the probability π_{BAY}, namely the probability of our subjects acting as bayesian in each of their decisions.
We tried to estimate the ρ_{NAI} and ρ_{BAY}, i.e. the risk aversion coefficients for both the approaches, and π_{BAY} altogheter, but the likelihood function did not converge. Indeed, the likelihood functions of our models are very close. For this reason, we estimated this probability numerically, using a grid loop.
Subsequently we let π_{BAY} moving inside a grid (0, 1), to finally choose the value that maximizes the likelihood function.
A possible drawback of this numerical procedure is the fact that the π_{BAY} standard error cannot be estimated, as it is shown in Table 3.2. On the other hand, we can justify this statement by saying that our likelihood function is not function of it, given that it is just a product of probability.
This numerical computation demonstrates that our subjects have a strong tendency to behave as naïve, given the estimation result of π_{BAY} = 0.2.

VARIABLES	
Rho_NAI	0.764***
	(0.00398)
Rho_BAY	0.637***
	(0.0109)
Pi_BAY	0.2
	-
Obs.	195,459

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3..2: Mixture Model with π_{BAY}=0.2

4. Conclusions

The thesis aims to explore how subjects approach ambiguous decision making when uncertainty is on a different level than the one usually investigated. In fact, although most decision problems in daily life involve different dimensions of uncertainty, the majority of models discussed in the economic literature deal with a specific form of uncertainty, the one on probability of the payoffs. In this investigation, uncertainty lies on the prizes of the lotteries people have to choose, rather than on their probabilities.

We explore the question of whether there is some systematic different behavior that people exhibit while dealing with ambiguity. Specifically, we investigate how a change in the information given to decision makers influences their risk aversion and, in order to find an answer to this question, we analyze the effects of some between subjects treatment.

Our results suggest that increasing the amount of ambiguity, people modify their predisposition toward it, showing some degree of love in favour of it. Namely, the lower the information given, the lower their risk aversion.

\\When the decision problem is faced under uncertainty, two different specifications are presented, the naïve and the bayesian. Indeed, according to the former, a pointwise estimation of the uncertain parameters is plugged inside the utility function. On the other hand, according to a more sophisticated paradigm, the bayesian, the prior distributions of these unknown parameters are used in order to compute the expected utility function, for each of them.

A mixture model demonstrates a strong majority of people, almost 80%, adopting the naïve approach.

In this thesis, we measured the treatment effects in the variation of ρ parameter in two models, namely ρ_{NAI} and ρ_{BAY}. Albarrán et al [1] adopt a different identification strategy, where the ρ is extracted from the TR2 subjects and the treatment effects are measured as the X and Y estimations. Although ours is a different approach, the results are more than compatible, that further confirm our main findings.

It is my intention, in the future, to apply more complex models of decision making under uncertainty to these data, like the models suggested by Klibanoff et al. [11].

This project, which constitutes the core of the future research advisable in this thesis, shall also incorporate the possibility to extend uncertainty to more and different levels, like already been done in Eliaz et al. [19]. Furthermore, we wonder whether individuals have a higher predisposition toward the bayesian approach while facing multiple levels of ambiguity.

This thesis, therefore, can be considered not only as a partial and exploratory analysis, but also as a good starting point for numerous and extended future investigations.

Bibliograpy

[1] Giovanni Ponti, Carmen Herrero, and Judit Albarràn, Pedro aan Alonso. An experiment on outcome uncertainty. Departamento de Fundamentos del Analisis Economico - Univesidad de Alicante, Working Paper, 2018.

[2] Charles A Holt and Susan K Laury. Risk aversion and incentive effects. American economic review, 92(5): 1644–1655, 2002.

[3] John D Hey and Chris Orme. Investigating generalizations of expected utility theory using experimental data. Econometrica: Journal of the Econometric Society, 62 (6): 1291–1326, 1994.

[4] Steffen Andersen, John Fountain, Glenn Harrison, E Elisabet Rutström, Estimating aversion to uncertainty. Department of Economics - Copenhagen Business School, Working Paper n° 7, 2009.

[5] Frank H Knight. Risk, uncertainty and profit, reprints of economic classics, Augustus M. Kelley, Bookseller, New York, 1964.

[6] Chip Heath and Amos Tversky. Preference and belief: Ambiguity and competence in choice under uncertainty. Journal of risk and uncertainty, 4(1): 5–28, 1991.

[7] Leonard J Savage. The foundations of statistics. Courier Corporation, North Chelmsford, MA, 1972.

[8] Daniel Ellsberg. Risk, ambiguity, and the savage axioms. The quarterly journal of economics, 75(4): 643–669, 1961.

[9] Barbara E Kahn and Rakesh K Sarin. Modeling ambiguity in decisions under uncertainty. Journal of Consumer Research, 15(2): 265–272, 1988.

[10] Marco Lauriola and Irwin P Levin. Relating individual differences in attitude toward ambiguity to risky choices. Journal of Behavioral Decision Making, 14(2): 107–122, 2001.

[11] Peter Klibanoff, Massimo Marinacci, and Sujoy Mukerji. A smooth model of decision making under ambiguity. Econometrica: Journal of the Econometric Society, 73(6): 1849–1892, 2005.

[12] Yoram Halevy. Ellsberg revisited: An experimental study. Econometrica: Journal of the Econometric Society, 75(2): 503–536, 2007.

[13] Larry G Epstein. A definition of uncertainty aversion. The Review of Economic Studies, 66(3): 579–608, 1999.

[14] Glenn W Harrison and E Elisabet Rutström. Expected utility theory and prospect theory: One wedding and a decent funeral. Experimental Economics, 12(2): 133, 2009.

[15] Daniel Kahneman. Prospect theory: An analysis of decisions under risk. Econometrica: Journal of the Econometric Society, 47(2): 278, 1979.

[16] Rakesh K Sarin and Robert L Winkler. Ambiguity and decision modeling: A preference- based approach. Journal of Risk and Uncertainty, 5(4): 389–407, 1992.

[17] Itzhak Gilboa and David Schmeidler. Maxmin expected utility with non-unique prior. Journal of mathematical economics, 18(2): 141–153, 1989.

[18] Carmen Herrero and Antonio Villar. Ranking distributions of monotone attributes. Department of Economics - Universidad Pablo de Olavide, Working Paper n° 14.03, 2014

[19] Kfir Eliaz and Pietro Ortoleva. Multidimensional ellsberg. Management Science, 62(8): 2179–2197, 2015.

[20] Jürgen Eichberger, Jörg Oechssler, and Wendelin Schnedler. How do subjects view multiple sources of ambiguity? Theory and Decision, 78(3): 339–356, 2015.

[21] Evan Moore and Catherine Eckel. Measuring ambiguity aversion. Unpublished manuscript, Department of Economics - Virginia Tech, 2003.

[22] Uri Gneezy, Alex Imas, and John List. Estimating individual ambiguity aversion: A simple approach. National Bureau of Economic Research, Working Paper n° 20982, 2015.

[23] Glenn W Harrison and E Elisabet Rutström. Risk aversion in the laboratory. In Risk aversion in experiments (Research in Experimental Economics, Vol. 12), pages 41–196, ed. by J. Cox and G. Harrison. Emerald Group Publishing Limited, Bingley, 2008.

[24] Hans P Binswanger. Attitudes toward risk: Experimental measurement in rural india. American journal of agricultural economics, 62(3): 395–407, 1980.

[25] Peter Wakker and Daniel Deneffe. Eliciting von neumann-morgenstern utilities when probabilities are distorted or unknown. Management science, 42(8): 1131–1150, 1996.

[26] Ben Greiner. An online recruitment system for economic experiments. In Forschung und Wissenschaftliches Rechnen pages 79–93, ed. by K. Kremer and V. Macho. GWDG Bericht Vol. 63. Gesellschaft für Wissenschaftliche Datenverarbeitung Göttingen: Datenverarbeitung, 2004.

[27] Shane Frederick. Cognitive reflection and decision making. Journal of Economic perspectives, 19(4): 25–42, 2005.

[28] Oliver P John, Eileen M Donahue, and Robert L Kentle. The big five inventory—versions 4a and 54. Berkeley, CA: Univeristy of California, Berkeley, Institute of Personality and Social Research, 1991

[29] Glenn W Harrison. Maximum likelihood estimation of utility functions using stata. Department of Economics, College of Business Administration - University of Central Florida, Working Paper n° 06–12, 2008.

[30] Peter G Moffatt. Experimetrics: Econometrics for experimental economics. Palgrave Macmillan, London, England, 2015.

[31] Ismael Rodriguez-Lara and Giovanni Ponti. Social motives vs social influence: An exper- iment on interdependent time preferences. Games and Economic Behavior, 105: 177–194, 2017.

[32] Graham Loomes and Robert Sugden. Testing different stochastic specificationsof risky choice. Economica, 65(260): 581–598, 1998.

[33] Gustav Fechner. Elements of psychophysics, Vol. 1. Holt, Rinehart and Winston, New York, NY, 1966

[34] Graham Loomes, Peter G Moffatt, and Robert Sugden. A microeconometric test of alternative stochastic theories of risky choice. Journal of risk and Uncertainty, 24(2): 103–130, 2002.

[35] John W Pratt, Howard Raiffa, and Robert Schlaifer. The foundations of decision under uncertainty: An elementary exposition. Journal of the American Statistical Association, 59(306): 353–375, 1964.

[36] Mark J Machina. Choice under uncertainty: Problems solved and unsolved. Journal of Economic Perspectives, 1(1): 121–154, 1987.

[37] Marco R Steenbergen. Maximum likelihood programming in stata. Department of Political Science. Department of Political Science - University of North Carolina, Chapel Hill, CA 2003.

[38] Judit Alonso. Risk aversion and uncertain outcomes in the laboratory. Departamento de Fundamentos del Analisis Economico - Univesidad de Alicante, Master Thesis, 2016.