CSEF
Centre for Studies in Economics and Finance

WORKING PAPER NO. 512

Unobserved Heterogeneity in Structural Behavioral Models Using Experimental Data

Giovanni Ponti and Claudio Rossetti

October 2018

University of Naples Federico II   **University of Salerno**   **Bocconi University, Milan**

WORKING PAPER NO. 512

Unobserved Heterogeneity in Structural Behavioral Models Using Experimental Data

Giovanni Ponti[*] and Claudio Rossetti[**]

Abstract

In this paper we compare a mixed logit model (MLM) and a latent class model (LCM) in the context of behavioral structural estimation using experimental data. By providing an instrument to deal with the intrinsic unobserved heterogeneity that characterizes experimental data, these alternative models have clear advantages compared with a multinomial logit model (MNL) typically used in structural estimation of behavioral models. We carry out our exercise by using experimental data that allows us estimation of distributional parameters related to risk and social preferences. Somehow coherently with the economic theory, the LCM identifies three classes of subjects (risk/ineq. lovers, risk/ineq. neutral, risk/ineq. averse). Moreover, estimates from both MLM and LCM somehow confirm the findings from a MNL model, that under the veil of ignorance (VOI) subjects' variance aversion mostly reflects risk, rather than distributional concerns. By taking unobserved heterogeneity adequately into account in the estimation of our structural behavioral model, also provides new insights into individual behavior on the interplay between risk and inequality concerns. For example, we find that there is much more variability in individual behavior when subjects face pure inequality than under VOI. Moreover, in the case of pure inequality subjects are also more likely to be inequality lovers than under VOI.

JEL Classification: D86; C25.

Keywords: Unobserved heterogeneity, Structural behavioral models, Social preferences, Risk preferences.

---

[*]   Universidad de Alicante, University of Chicago, LUISS Guido Carli Roma.

[**] Università di Napoli Federico II and CSEF. E-mail: claudio.rossetti@unina.it

Table of contents

1 Introduction

Structural estimation of behavioral models using experimental data is becoming a standard in the analysis of many decision contexts, such as risk and uncertainty (Harrison and Rustrom [19]), intertemporal decisions (Harrison *et al.* [1]) or distributional choices (Fisman *et al* [13]). This literature has the great advantage of reducing the gap between theoretical and empirical aspects of behavior in that the experimental evidence is framed directly within the models that have been proposed to explain it. Powerful maximum-likelihood estimation packages (now available within every standard statistical software) allow scholars (not only experienced econometricians) to estimate the parameters of classic models of decision, or validating new ones.

The most popular exercise of structural estimation consists in framing the (discrete) choice problem using the multinomial logit model (MNL, see Harrison [17]). More recently, few papers have corrected some basic drawback of the MNL model by allowing for between-subjects heterogeneity by way of two alternative techniques. The first technique consists of estimating *individual fixed effects* (FE, see, e.g., Cabrales *et al.* [6] and Frignani and Ponti [14]), which – essentially – yields the estimation of fixed coefficients for each subject participating to the experiment. The second technique consists of applying the so-called mixed logit model (MLM, see, e.g., Bellemare et al. [3], or Conte et al. [8]), which -essentially-yields the estimation of random coefficients characterized by an individual component drawn from a (parametrically imposed) common distribution.

These alternative approaches have clear advantages compared with MNL in the context of structural estimation, since they provide an instrument to deal with the intrinsic unobserved heterogeneity that characterizes experimental data. Moreover, this observation has prompted the theoretical discussion in Behavioral Economics, by introducing models in which players are characterized by "behavioral types" - take for instance the partition of social preferences among "inequity averse", "status seekers" , "efficiency seekers", or "egoists" in Dictator Games (Engelmann and Strobel [9]), or the partition between "un/conditional cooperators" and "selfish" in he context of (repeated) public good provision (Fischbacher and Gachter [12]). These models do not treat heterogeneity as associated with individuals' observable characteristics - such as gender, or income - but measure it as preference parameter realizations revealed by subjects' actual behavior.

In this paper we compare the MLM technique with an alternative approach, namely the

1

latent class model (LCM), in the context of behavioral structural estimation in order to analyze the sources of the individual heterogeneity. We do so within the realm of a simple (mean-variance) random utility specification. LCMs have been extremely popular especially in psychology and marketing. In this approach, each latent class consists of a number of individuals that are assumed to be homogeneous with respect to their preferences for alternatives. Latent classes, however, differ in preference, meaning that the taste parameters differ between latent classes. This model can be very useful to read evidence of many experiments, where subjects' behavior is framed within the realm of a specific random utility model, where the partition in specific subsets of the relevant parameter space has a natural correspondence with a "behavioral type", within the realm of the behavioral model under scrutiny.

To carry out our exercise we will use experimental data from Cabrales et al. [6] (CABRA hereafter) and Frignani and Ponti [14] (FRIGNO hereafter). In these two papers, distributional preference parameters are estimated by way of a simple Dictator Game in which subjects repeatedly select their favorite option among a fixed menu of four, which changes at every round. Each option consists in two monetary prizes, one for them, one for another (randomly and anonymously matched) subject participating to the experiment. The exercise carried out there exploits the experimental methodology by designing specific economic environments in which, sometimes risk and inequality concerns are isolated, sometimes they are combined by facing inequality under the veil of ignorance (VOI), checking how these alternative decision frames affect the estimates of the same parameter, under the same statistical model.

The comparison of the estimated parameters across different environments detects how subjects react to (absolute) difference in payoff. In this respect, FRIGNO's main result -namely, positive and significant variance aversion in all environments, and (no) significant difference comparing the parameter associated with pure inequality aversion (pure risk aversion) compared with the control VOI - yields a natural interpretation: when facing choices under the VOI, subjects seem to undermine the "inequality" dimension of the problem (which is, instead significant in absence of risk), and their choices are not distinguishable from those taken in a simple lottery environment.

Nevertheless, taking unobserved heterogeneity in risk and social preferences adequately into account turns out to be a critical issue in the estimation of such a structural behavioral

model. FRIGNO's empirical conclusions are derived from a MNL model, i.e., without controlling for individual unobserved heterogeneity. First of all, somehow coherently with the economic theory, the LCM identifies three classes of subjects (risk/ineq. lovers, risk/ineq. neutral, risk/ineq. averse). Moreover, estimates from both MLM and LCM somehow confirm the findings in Frignani and Ponti [14], derived from a MNL model, that under VOI subjects' variance aversion mostly reflects risk, rather than distributional concerns. By taking unobserved heterogeneity adequately into account in the estimation of our structural behavioral model also provides new insights into individual behavior about risk and inequality. For example, we find that there is much more variability in individual behavior when subjects face pure inequality than under VOI. Moreover, in the case of pure inequality subjects are also more likely to be inequality lovers than under VOI.

The remainder of this paper is arranged as follows. In Section 2 we briefly describe the experimental design. Section 3 describes our empirical model of preferences. Section 4 presents our results. Finally, Section 5 offers some conclusions.

2   Experimental design

In what follows, we describe the features of FRIGNO's experimental environment. Starting from CABRA, where player position (i.e. the identity of the best paid agent) is constant across options, and known in advance before subjects have to make their decisions (treatment T_1), FRIGNO complement their evidence with two additional treatments in which subjects face the same sequence of decisions under VOI, knowing ex-ante that either player position is equally likely (T_2), or the same sequence of decisions is made under a "lottery frame", in which player position is unknown (and equally likely), but there is no payoff externality on others (T_3). Here the experimental evidence is read by the way of a simple mean/variance utility model where, depending on the treatments, the parameter associated with the variance measures pure inequality aversion (T_1), pure risk aversion (T_3) or, in the VOI treatment T_2, some combination of the two.

2.1   Sessions

They run 8 sessions under 3 different treatments: 3 sessions for each of the "Dictator Game" treatments, T_1 and T_2, 2 sessions for the "Lottery" treatment, T_3. All experimental sessions

were conducted at the Laboratory of Theoretical and Experimental Economics (LaTEx), of the Universidad de Alicante. A total of 192 students (24 per session) were recruited among the undergraduate population of the Universidad de Alicante. The experimental sessions were computerized.[1] Instructions were read aloud and we let subjects ask about any doubt they may have had. In all sessions (but those of the lottery treatment T_3), subjects were divided into two *matching groups* of 12, with subjects from different matching groups never interacting with each other throughout the session.[2]

All treatments share the same basic layout. At the beginning of each round $t = 1, \ldots, 24$, subjects are informed about the choice set $C_t = \{a_{kt} : k = 1, \ldots, 4\}$, constant across treatments. Each option $a_{kt} = (a_{kt}^1, a_{kt}^2)$ assigns a fixed monetary prize, a_{kt}^j, to player $j = 1, 2$, with $a_{kt}^1 \geq a_{kt}^2$, $\forall k$. After subjects have selected their favorite options, all payoff relevant information is revealed and payoffs are distributed.

2.2 Treatments

We now explain the details of our 3 experimental conditions.

T_1 : pure inequality (CABRA)

In T_1, subjects choose their preferred option after being informed about the outcome of the iid draw which fixes the player (i.e. their relative) position for that pair and round. Remember that, since $a_{kt}^1 \geq a_{kt}^2$, player 1 (2) looks at the distributional problem implicit in the option choice from the viewpoint of the (dis)advantaged player. Once choices are made, we employ a "Random Dictator" protocol (Harrison and McDaniel [18]) to determine the payoff relevant decision, in that another iid draw fixes the identity of the subject whose choice determines the monetary rewards for that pair and round.

T_2 : VOI

In treatment T_2 we modify the control treatment T_1 by introducing the VOI. In this case, subjects only know that the ex-ante probability of being assigned to either player position

---

[1] The experiment was programmed and conducted with the software z-Tree (Fischbacher [11]). The complete set of instructions, translated into English, can be found in the Appendix.

[2] Given this design feature, we shall read the data under the assumption that the history of each matching group corresponds to an independent observation. Clearly, the same does not apply in case of T_3, in that each subject's experimental history corresponds to an independent observation.

equals $\frac{1}{2}$. Everything else is just as in T_1, in particular the fact that subjects alternate player and Dictator positions in a (iid) random fashion.

T_3 : pure risk (LOTT)

Our lottery treatment T_3 replicates treatment T_2 without any payoff externality on others. In this case, player position is uncertain (and equally likely), but each subject decides in isolation.

2.3 Financial rewards

All monetary payoffs in the experiment were expressed in Spanish Pesetas (1 euro is approx. 166 ptas.). Subjects received 1.000 ptas. just to show up, to which they summed up all their cumulative earnings throughout the 24 rounds of the experiment. Average earnings were about 12 euros, for an experimental session lasting for approximately 45 minutes.

3 An empirical model of preferences

Subject i faces a choice among the alternatives in choice set C_t in each of the 24 choice situations (rounds). We model the utility that subject $i = 1, \ldots, n$ obtains from alternative $k = 1, \ldots, 4$ in choice occasion $t = 1, \ldots, 24$ as

$$V_{ikt} = \mu_{ikt} - \gamma_i s_{ikt},$$

where μ_{ikt} is either the monetary payoff assigned to the player in T_1 or the average of the two players' payoffs otherwise, and s_{ikt} is a measure of the distance between the two players' payoffs. Specifically, we model s_{ikt} as the standard deviation of the two players' payoffs for option k in choice occasion t. Thus, our model corresponds to a simple mean-variance utility model.

The unobserved coefficient γ_i is allowed to vary over subjects and its interpretation depends on the treatment.[3]. In treatment T_1 subjects are allowed to care about both her own payoffs and the payoffs of the other player and they are informed about their player position before they choose. Thus, in this case, we can think of γ_i as a direct measure of *pure inequality aversion* related to a pure distributional concern. For subjects who only care

---

[3] Note that the model can be generalized to allow γ_i to vary also over rounds, t

about their own payoff, γ_i would be zero. In treatment T_3 subjects face ordinary binary lotteries. Therefore, in this case, γ_i measures pure risk aversion, with $\gamma_i = 0$ indicating the null hypothesis of risk-neutrality. In T_2 subjects chooses among different lotteries, but this decision has distributional consequences for both players. In this sense, we expect γ_i to capture some combination of both effects, one related to risk, the other to inequality.

Perfect optimization would imply that, in choice occasion t, subject i chooses the alternative k that maximizes her expected utility V_{ikt}. To allow for suboptimal choices, we add idiosyncratic error terms $\lambda_i \epsilon_{ikt}$ and we assume that subject i chooses the alternative k that maximizes

$$U_{ikt} = V_{ikt} + \lambda_i \epsilon_{ikt},$$

where the size of the noise parameter λ_i drives the likelihood of suboptimal choice. We also assume that the errors ϵ_{ikt} are distributed iid extreme value over subjects, alternatives and rounds, independent of each other and the other variables in the model. This is equivalent to assume that the difference of any two ϵ_{ikt} across alternatives follows a logistic distribution. The probability that individual i chooses option j at round t is therefore standard logit

$$P_{ijt}(\gamma_i) = \frac{\exp(V_{ijt}/\lambda_i)}{\sum_{k=1}^{4} \exp(V_{ikt}/\lambda_i)}.$$

Let j_{it} denote the alternative that subject i chose at round t. Since ϵ_{ikt} are independent over choice occasions, the probability of subject i's observed sequence of choices is

$$P_i(\gamma_i) = \prod_{t=1}^{24} P_{ij_{it}t}(\gamma_i).$$

Assuming that γ_i has a distribution G with a known shape, the unconditional probability for the observed sequence of choices is computed by taking the expectation of $P_i(\gamma_i)$ with respect to γ_i

$$P_i = \int P_i(\gamma_i) \, dG(\gamma_i).$$

Thus we have a mixture of the logit function evaluated at different γ_i's with the density of γ_i as the mixing distribution.

Let θ be the model parameters. Given a probabilistic sample of n subjects, each observed for 24 rounds, a maximum likelihood (ML) estimator of θ is obtained by maximizing the sample log-likelihood

$$\mathcal{L}(\theta) = \sum_{i=1}^{n} \ln P_i(\theta). \tag{1}$$

6

The standard MNL is a special case where the mixing distribution is degenerate at fixed parameters. Moreover, this specification does not exhibit IIA and its restrictive substitution patterns implied by a standard MNL. In fact, the ratio of any two logit probabilities depends on all the data, including attributes of other alternatives. The mixing distribution of γ_i can be either discrete or continuous. In the former case we have a LCM, while in the latter case we have a MLM.

3.1 MLM

Under MLM the mixing distribution of γ_i is assumed to be continuous. In this case, it is assumed that individual specific parameters γ_i are continuously distributed across individuals, and they are modeled as

$$\gamma_i = X_i'\beta + \sigma u_i,$$

where β and σ are unknown parameters, X_i is a vector consisting of the intercept and observed individual characteristics, and u_i is an unobserved individual effect.

The term u_i reflect unobserved heterogeneity and is assumed to be distributed independently of the other variables and the error term in the model, with zero mean, unit variance and a distribution function G with a known shape. Any continuous distribution such as normal, lognormal, uniform, triangular, gamma, etc. can be actually used. In this case, the choice probability is

$$P_{ikt}(\gamma_i) = \frac{\exp[(\mu_{ikt} - \gamma_i s_{ikt})/\lambda_i]}{\sum_{k=1}^4 \exp[(\mu_{ikt} - \gamma_i s_{ikt})/\lambda_i]}.$$

Let θ denote the vector of parameters to be estimated. The resulting choice probability is the expectation of the conditional probability $P_i|u_i$ with respect to u_i

$$P_i(\theta) = \int P_i|u_i \, dG(u_i).$$

In this case, exact maximum likelihood estimation is typically not feasible, because this integral does not generally have a closed form solution. Nonetheless, it can be approximated by simulation. The simulated probability is the sample average

$$\tilde{P}_i(\theta) = \frac{1}{R} \sum_{r=1}^R P_i|u_i^r, \tag{2}$$

where $\{u_i^r : r = 1, \ldots, R\}$ are R independent draws from the distribution G.

Thus, estimating the parameter vector θ by maximum simulated likelihood (MSL) involves maximizing an approximation to $\mathcal{L}(\theta)$, obtained by replacing the probability $P_i(\theta)$ with simulated probabilities $\tilde{P}_i(\theta)$. For a discussion of simulation-based estimators, see Hajivassiliou and Ruud [16]. The simulated log-likelihood is then

$$\tilde{\mathcal{L}}(\theta) = \sum_{i=1}^{n} \ln \tilde{P}_i(\theta).$$

If the number of draws R increases with the sample size n, MSL provides an estimator of the parameters θ which is consistent as $n \to \infty$, asymptotically normal, and asymptotically equivalent to the ML estimator (Lee [21], Hajivassiliou and Ruud [16], Gouriéroux and Monfort [15]).

The elements of the vector u_i^r are independently drawn from the assumed distribution G. In practice, instead of using pseudo-random draws to obtain u_i^r, we base the simulation on Halton sequences (Train [25]). Halton sequences generate draws that provide a more systematic coverage of the domain of integration than independent random draws. They usually imply a lower integration error and faster convergence rates, and require a smaller number of draws. Since Halton sequences are deterministic, following Wang and Hickernell [26] we introduce randomness by using a random start procedure. Specifically, we draw an integer randomly between 0 and some large K and label the draw N_0. Then, we create a Halton sequence starting at integer N_0 in step 1 above.

Using Bayes theorem, we can obtain posterior estimates of the individual-specific parameter vector

$$\hat{\gamma}_i = \frac{\frac{1}{R} \sum_{r=1}^{R} (\hat{P}_i | u_i^r \; \hat{\gamma}_i^r)}{\frac{1}{R} \sum_{r=1}^{R} \hat{P}_i | u_i^r}. \tag{3}$$

3.2 LCM

Under LCM, it is assumed that individuals are sorted into a set of M classes, but which class contains any particular individual, whether known or not to that individual, is unknown to the analyst. In other words, for each subject, γ_i can take M possible values $\gamma_1, \ldots, \gamma_M$ with prior probability H_{im} that subject i is in class m (i.e. $\gamma_i = \gamma_m$). In this case, the choice probability is

$$P_{ikt}(\gamma_m) = \frac{\exp[(\mu_{ikt} - \gamma_m s_{ikt})/\lambda_i]}{\sum_{k=1}^{4} \exp[(\mu_{ikt} - \gamma_m s_{ikt})/\lambda_i]}.$$

This specification is particularly useful if one believes that there are M segments in the population, each with its own choice behavior or preferences. Consider, for example, the population consists of individuals who are risk adverse, risk neutral or risk lovers. The share of the population in each segment can be estimated within the model.

Given that the class assignment is unknown, we can conveniently model the prior probability for class m for subject i as

$$H_{im} = \frac{\exp(Z_i'\delta_m)}{\sum_{m=1}^{M} \exp(Z_i'\delta_m)},$$

where Z_i is a vector consisting of the intercept and a set of observable individual characteristics which enter the model for class membership, and the parameter vector δ_M is normalized to zero to secure identification of the model.

Let θ denote the vector of parameters to be estimated (consisting of the M structural parameter vectors γ_m and the $M-1$ latent class parameter vectors δ_m). The resulting choice probability is the expectation (over classes) of the class-specific probabilities

$$P_i(\theta) = \sum_{m=1}^{M} H_{im} P_{i|m}.$$

An issue is the choice of the number of classes M. Since M is not a parameter in the interior of a convex parameter space, one cannot test hypotheses about M directly. One possibility is hence to use the Bayesian information criterion (BIC) to choose M.

Even in this case, we can obtain posterior estimates of the individual-specific parameter vector

$$\hat{\gamma}_i = \sum_{m=1}^{M} \hat{H}_{m|i}\hat{\gamma}_m, \tag{4}$$

where

$$\hat{H}_{m|i} = \frac{\hat{P}_{i/m}\hat{H}_{im}}{\sum_{m=1}^{M} \hat{P}_{i/m}\hat{H}_{im}}$$

is a posterior estimate of the latent class probabilities (conditioned on estimated choice probabilities). The class associated with the maximum value of $\hat{H}_{m|i}$ would be a strictly empirical estimator of the latent class within which the individual is.

4 Results

In our empirical specification we assume normality of the individual effects u_i in the MLM, and we set the noise parameter $\lambda_i = 1$. Results of a MLM where the unobserved individual effect u_i is assumed to have a symmetric Triangular distribution are reported in Appendix A. Moreover, MLM parametrizes the mean and the standard deviation of the underlying distributions to differ among treatments. Three latent classes are selected in the LCM as the best fit from 2, 3, 4 and 5 classes. The LCM assumes a fixed parameter vector in each class, with the overall mean in each treatment being a function of how these are mixed by the class probabilities.

Table 1 shows estimated coefficients of the MNL, MLM and LCM. Note that evaluating the absolute parameter estimates across models is not informative because of scale differences. It will be much more meaningful to base any comparison on model predictions such as choice probabilities, elasticities or simulations. Nevertheless, comparing the absolute parameters estimates is beyond the aim of these paper. Instead, we are interested here in what different models can tell us about how estimated parameters differ across treatments.

As Table 1 shows, controlling for between-subject heterogeneity has a critical impact on the conclusions of the model. Frignani and Ponti's [14] MNL estimates are reported in the first column of Table 1. Here we see that, when we neglect heterogeneity, all three coefficients turn out to be positive and highly significant, indicating high "variance aversion" in all choice domains. As for their relative comparison, we note that variance aversion in T_1 is significantly smaller than in the control treatment, T_2, while we cannot reject the null that γ in T_2 and T_3 are equal. FRIGNO interpret this evidence as $t$he predominance of risk (rather than distributional) concerns under the VOI. Things change when we look at the MLM estimates (Table 1, second column). Here we see that the estimated coefficient in T_1 is the highest, while variance aversion in T_3 is smaller than in the other two treatments. Estimated standard deviation is also higher in T_1 (Table 1, second column and second block of rows), denoting higher variability in the individual-specific parameters.

We now move to the analysis of the LCM estimates (Table 1, last three columns). Here we find that, for T_2 and T_3, our three-class partition nicely adapts to the mean-variance economic interpretation, in that the (left) central [right] class is characterized by a coefficient which is not significantly different (smaller) [greater] than zero (indicating $v$ariance neutrality (loving) [aversion],respectively). This in not the case for T_1, $w$hose estimated coefficients of classes 2

10

and 3 are both positive and significant, with Class 3 (the class with the highest coefficient) characterized by an estimated γ comparatively higher than the corresponding estimates in T_2 and T_3 (1.312 against .424 and .360, respectively). We also see that T_1's coefficient for Class 1 is also higher (compared with the other 2). In this sense, the entire coefficient distribution for T_1 is shifted on the right, compared to those of T_2 and T_3.

As for the estimated latent class probabilities (Table 1, last three columns and second block of rows), in the case of T_1 extreme classes have similar probabilities (.153 and .227, respectively). By contrast, estimated probabilities for T_2 and T_3 suggest a parameter distribution positively skewed (since, in both cases, class probabilities are increasing with the classes).

Table 2 shows posterior estimates of the risk/inequality aversion coefficients for MLM and LCM. Specifically, to characterize the sample, we average over individuals the posterior estimates of the individual-specific parameters (computed as in Equations (3) and (4) respectively). To judge their accuracy, posterior estimates are calculated 500 times, for 500 independent draws of the parameters from the estimated asymptotic distribution of their estimator. We present the median, and the fifth and ninety-fifth percentiles. The latter two are the bounds of a two sided confidence interval of approximately 90 percent. Here we see that MLM and LCM estimates of the average risk/inequality individual parameters are consistent to each other. Specifically, the order of the average risk/inequality parameters is the same between MLM and LCM, with the coefficient in T_3 ranking the lowest and the coefficient in T_1 the highest. The 90% confidence intervals of the parameters in T_1 and T_3 never overlap, while the confidence intervals of the parameters in T_1 and T_2 overlap only in the case of the MLM. Considering both MLM and LCM the average of the γ_i's over the individuals is not significantly different in the case of T_2 and T_3.

Figures 1 and 2 report a graphic sketch of the distribution of the posterior estimates of the individual risk/inequality aversion coefficients for MLM and LCM respectively. These Figures seem to confirm FRIGNO's finding that behavior in T_2 and T_3 is remarkably similar, and that a remarkably lower fraction of the estimated individual γ_i's is negative as compared to T_1. To stress this last point, Table 3 shows the partition of subjects pool with respect to their attitude toward risk/inequality (percentage in parenthesis). Partition is based on posterior estimates of the individual-specific risk/inequality aversion parameters. Posterior estimates are calculated 500 times, for 500 independent draws of the parameters from the

estimated asymptotic distribution of their estimator. From these 500 replications of the individual-specific parameters a subject is classified as risk/inequality lover if both the $5th$ and the $95th$ percentiles are negative, as risk/inequality averse if both the $5th$ and the $95th$ percentiles are positive, and as risk/inequality neutral otherwise. The percentages in T_2 and T_3 are very similar, with a few subjects classified as risk/inequality lover. In T_1 there are very few inequality neutral subjects, while the percentage of inequality lovers is remarkably higher than in the control treatment T_2.

Thus, if we look at the individual-specific risk/inequality parameters γ_i both MLM and LCM seem to somehow confirm FRIGNO's results that there is *a* predominance of risk, rather than distributional, concerns under the VOI. Nevertheless, by taking unobserved heterogeneity adequately into account in the estimation of our structural behavioral model also provides new insights into individual behavior about risk and inequality. For example, we find that there is much more variability in individual behavior when subjects face pure inequality without VOI (T_1) than in T_2 and T_3. Moreover, in the case of pure inequality subjects are also more likely to be inequality lovers than under VOI.

Finally, appendix A shows results of alternative specifications for the MLM and the LCM. We report estimates of a MLM where the unobserved individual effect u_i is assumed to have a symmetric Triangular distribution, and a LCM where the coefficients of the second class are constrained to zero (risk/inequality neutral). In the latter case, a likelihood ratio test rejects the hypothesis the all the coefficients of the second class are equal to zero (χ^2=43.3, p-value=0.000). This is mainly due to the coefficient in treatment 1. Specifically, we report estimated coefficients, posterior estimates of risk/inequality aversion, and the partition of subjects pool with respect to their attitude toward risk/inequality. Estimation of these models mainly confirms our previous findings.

5 Conclusion

In this paper we compare a mixed logit model and a latent class model in the context of behavioral structural estimation. These alternative models have clear advantages compared with a multinomial logit model in the context of structural estimation, since they provide an instrument to deal with the intrinsic unobserved heterogeneity that characterizes experimental data. We carry out our exercise within the realm of a simple (mean-variance) random utility specification by using experimental data from Cabrales et al. [6] and Frignani and

Ponti [14], where risk and inequality concerns are sometimes isolated, sometimes they are combined by facing inequality under the veil of ignorance, allowing us to analyze how alternative decision frames affect the estimates of the same parameter, under the same statistical model.

Overall, we find that both the LCM and the MLM represent a remarkable statistical improvement over the MNL. Hence, based on the empirical evidence herein, we believe that both LCM and MLM offer attractive specifications to model unobserved heterogeneity in individual preferences. Which model between LCM and MLM is superior is somehow inconclusive. Nonetheless, we believe that this is perhaps an encouraging result, motivating the researcher to compare different specifications of the choice process. Each model has its own pros and cons. On the one hand, the LCM has the advantage of being a semiparametric specification, avoiding possibly strong distributional assumptions about individual heterogeneity. On the other hand, the MLM, while fully parametric, can be sufficiently flexible that it provides wide range of possibilities to specify individual unobserved heterogeneity, and it is usually more parsimonious than LCM.

Specifically, in our application to risk and social preferences under VOI, somehow coherently with the economic theory, the LCM identifies three classes of subjects (risk/ineq. lovers, risk/ineq. neutral, risk/ineq. averse). Moreover, estimates from both MLM and LCM somehow confirm the findings in Frignani and Ponti [14], derived from a MNL model, that under VOI subjects' variance aversion mostly reflects risk, rather than distributional concerns. By taking unobserved heterogeneity adequately into account in the estimation of our structural behavioral model also provides new insights into individual behavior about risk and inequality. For example, we find that there is much more variability in individual behavior when subjects face pure inequality than under VOI. Moreover, in the case of pure inequality subjects are also more likely to be inequality lovers than under VOI.

References

[1] Andersen S, Harrison GW, Lau MI and Rutström E (2008). "Eliciting Risk and Time Preferences", *Econometrica* 76 , 583–618.

[2] Andreoni J and Miller J (2002). "Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism", *Econometrica*, 70, 737-753.

[3] Bellemare, C., Kröger, S., Van Soest, A. (2008). "Measuring Inequity Aversion in a Heterogeneous Population Using Experimental Decisions and Subjective Probabilities", *Econometrica*, 76(4), 815-39.

[4] Bolton G, Brandts J. and Ockenfels A. (2005). "Fair Procedures: Evidence from Games Involving Lotteries", *Economic Journal*, 115, 1054–1076.

[5] Bosmans K and Schokkaert E (2003). *Social Welfare, the Veil of Ignorance and Purely Individual Risk: An Empirical Examination*, mimeo.

[6] Cabrales A, Miniaci R, Piovesan M and Ponti G (2008), *Social Preferences and Strategic Uncertainty: an Experiment on Markets and Contracts*, *American Economic Review*, 100, 2261-78.

[7] Carlsson F , Daruvala F, Johansson-Stenman O (2005). "Are People Inequality-Averse, or Just Risk-Averse?", *Economica,* 72, 375–396.

[8] Conte, A., Hey J.D. and Moffatt, P.G. (2011). "Mixture models of choice under risk" *Journal of Econometrics*, 162, 79-82.

[9] Engelmann, D. and Strobel, M. (2004). "Inequality Aversion, Efficiency, and Maximin Preferences in Simple Distribution Experiments", *American Economic Review*, 94(4), 857–869.

[10] Fehr E and Schmidt KM (1999). "A theory of fairness,competition and cooperation",*Quarterly Journal of Economics*, 114, 817 –868.

[11] Fischbacher U (2007). "z-Tree: Zurich Toolbox for Ready-made Economic Experiments." *Experimental Economics,* 10(2): 171-178.

[12] Fischbacher, U. and Gächter, S. (2010). "Social Preferences, Beliefs, and the Dynamics of Free Riding in Public Goods Experiments", *American Economic Review,* 100(1), 541-556.

[13] Fisman R, Kariv S and Markovits D (2007). "Individual Preferences for Giving". *American Economic Review* 97(5): 1858-76.

[14] Frignani, N. and Ponti, G. (2012). "Social vs. Risk Preferences under the Veil of Ignorance", *Economics Letters*, 116, 143-6.

[15] Gouriéroux, C. and A. Monfort (1997), *Simulation Based Econometric Methods*, New York: Oxford University Press.

[16] Hajivassiliou, V.A. and P.A. Ruud (1994). *Classical estimation methods for ldv models using simulation*, in R.Engle and D.McFadden, eds, *Handbook of Econometrics*, Vol. 4, New-York: Elsevier, pp. 2383-2441.

[17] Harrison, G.W. (2007). Maximum Likelihood Estimation of Utility Functions Using Stata, Working Paper 06-12, Department of Economics, College of Business Administration, University of Central Florida

[18] Harrison GW and McDaniel T (2008). "Voting Games and Computational Complexity", *Oxford Economic Papers*, 60, 546-565.

[19] Harrison, G. and Rutström, E. (2009). "Expected Utility Theory and Prospect Theory: One Wedding and A Decent Funeral", *Experimental Economics* **12**, 133-158.

[20] Harsanyi, J. C. (1977), *Rational Behavior and Bargaining Equilibrium in Games and Social Situations*, Cambridge University Press, Cambridge.

[21] Lee, L. (1992). "On effciency of methods of simulated moments and maximum simulated likelihood estimation of discrete response models", *Econometric Theory* 8, 518-52.

[22] Rawls J (1971). *A Theory of Justice*, Cambridge, Mass., Harvard University Press.

[23] Rousseau JJ (1762). *The Social Contract*, London, Dent, 1973.

[24] Sobel J (2005). "Interdependent Preferences and Reciprocity," *Journal of Economic Literature,* 43, 392-436.

[25] Train, K. (2003), *Discrete Choice Methods with Simulation*, Cambridge, UK: Cambridge University Press.

[26] Wang, X. and Hickernell, F. (2000). *Randomized Halton sequences*, Mathematical and Computer Modeling, 32, 887–899.

Table 1: Estimated coefficients (* significant at 10%, ** significant at 5%, *** significant at 1%).

	MNL	MLM	LCM Class 1	Class 2	Class 3
T1	0.267 ***	0.288 ***	-0.595 ***	0.189 ***	1.312 ***
T2	0.347 ***	0.244 ***	-1.107 ***	0.064	0.424 ***
T3	0.316 ***	0.150 *	-1.080 ***	-0.071	0.360 ***
		σ	Latent class probability		
T1		0.683 ***	0.153 **	0.620 ***	0.227 ***
T2		0.401 ***	0.032	0.436 ***	0.531 ***
T3		0.327 ***	0.071 *	0.279 **	0.650 ***
Log-like.	-6,369.2	-5,934.4	-5,978.2		
BIC	13,087.4	12,377.5	12,492.5		

Notes: 192 subjects, 24 rounds, 4 alternatives.

Table 2: Posterior estimates of risk/inequality aversion.

	MLM P5	Median	P95	LCM P5	Median	P95
T1	0.271	0.277	0.286	0.304	0.325	0.340
T2	0.229	0.251	0.284	0.155	0.214	0.286
T3	0.091	0.148	0.242	0.100	0.141	0.180

Notes: P5: $5th$ percentile; P95: $95th$ percentile.

Table 3: Partition of subjects pool with respect to their attitude toward risk/inequality (percentage in parenthesis).

	MLM Lover	Neutral	Averse	LCM Lover	Neutral	Averse
T1	23	4	45	10	6	56
	(31.9)	(5.6)	(62.5)	(13.9)	(8.3)	(77.8)
T2	5	21	46	3	24	45
	(6.9)	(29.2)	(63.9)	(4.2)	(33.3)	(62.5)
T3	3	18	27	4	14	30
	(6.2)	(37.5)	(56.2)	(8.3)	(29.2)	(62.5)
Tot	31	43	118	17	44	131
	(16.1)	(22.4)	(61.5)	(8.9)	(22.9)	(68.2)

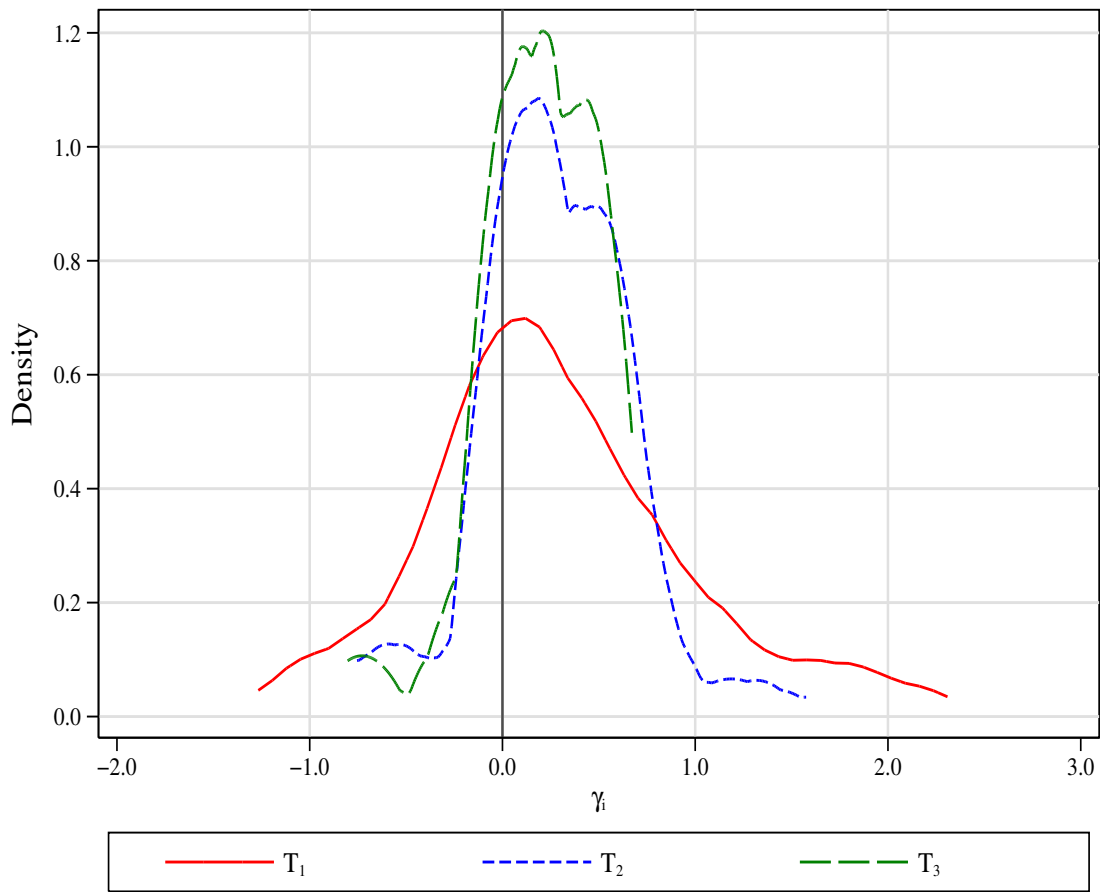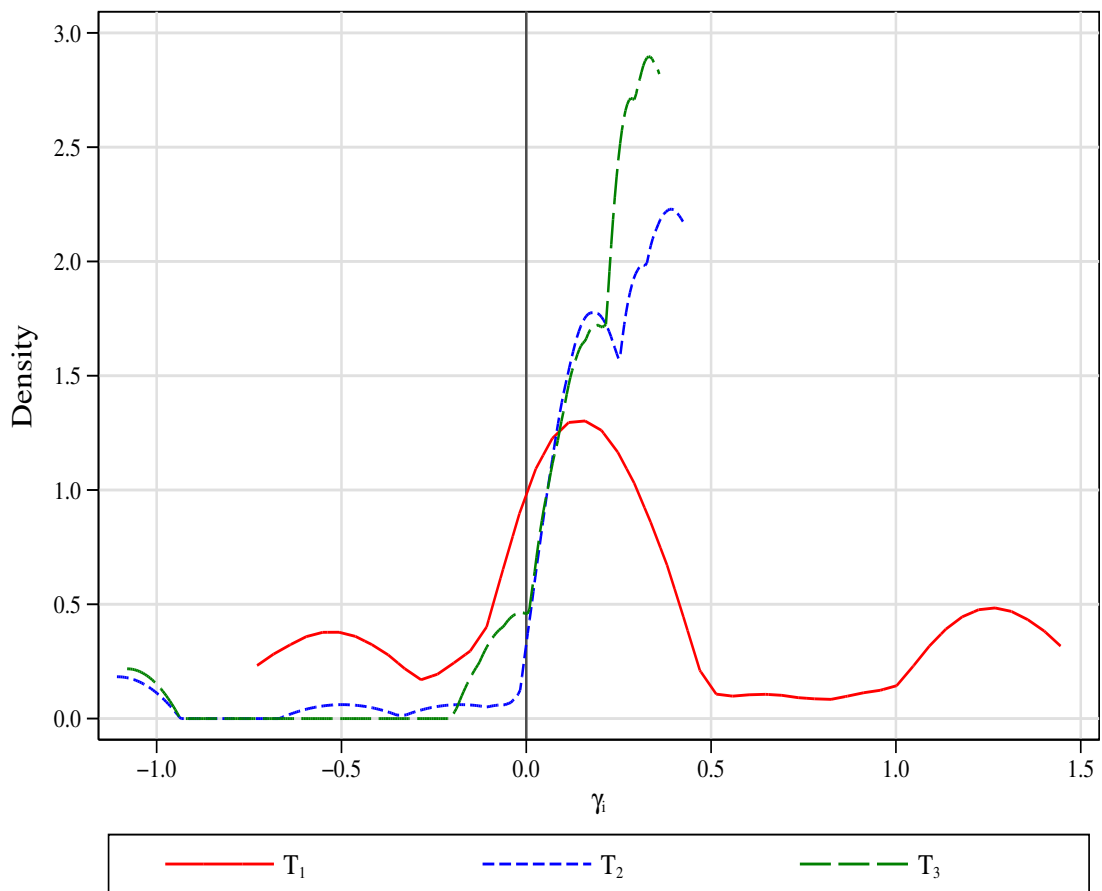Figure 1: Kernel density estimate for γ_i (MLM).

Figure 2: Kernel density estimate for γ_i (LCM).

Appendix

A Alternative specifications for the MLM and the LCM

This Appendix presents results of alternative specifications for the MLM and the LCM. We report estimates of a MLM where the unobserved individual effect u_i is assumed to have a symmetric Triangular distribution, and a LCM where the coefficients of the second class are constrained to zero (risk/inequality neutral).

Estimated coefficients (* significant at 10%, ** significant at 5%, *** significant at 1%).

	MLM	LCM		
		Class 1	Class 2	Class 3
T1	0.386 ***	-0.644 ***	0.000	1.287 ***
T2	0.278 ***	-1.126 ***	0.000	0.420 ***
T3	0.115	-1.064 ***	0.000	0.363 ***
	Scale	Latent class probability		
T1	1.983 ***	0.098 **	0.648 ***	0.254 ***
T2	1.153 ***	0.031	0.422 ***	0.547 ***
T3	0.774 *	0.074 *	0.297 ***	0.629 ***
Log-like.	-5,947.4		-5,999.8	
BIC	11,958.0		12,125.9	

Posterior estimates of risk/inequality aversion.

	MLM			LCM		
	P5	Median	P95	P5	Median	P95
T1	0.270	0.279	0.285	0.253	0.263	0.272
T2	0.221	0.245	0.280	0.185	0.195	0.205
T3	0.068	0.148	0.312	0.125	0.148	0.176

Notes: P5: 5th percentile; P95: 95th percentile.

Partition of subjects pool with respect to their attitude toward risk/inequality (% in parenthesis).

	MLM			LCM		
	Lover	Neutral	Averse	Lover	Neutral	Averse
T1	24	2	46	41	3	28
	(33.3)	(2.8)	(63.9)	(56.9)	(4.2)	(38.9)
T2	10	17	45	3	9	60
	(13.9)	(23.6)	(62.5)	(4.2)	(12.5)	(83.3)
T3	0	21	27	4	4	40
	(0)	(43.8)	(56.2)	(8.3)	(8.3)	(83.3)
Tot	34	40	118	48	16	128
	(17.7)	(20.8)	(61.5)	(25.0)	(8.3)	(66.7)