Netspar Network for Studies on Pensions, Aging and Retirement

Hans-Martin von Gaudecker

Arthur van Soest

Erik Wengström

Risk Preferences in the Small for a Large Population

Discussion Paper 2008 - 011

April 18, 2008

Risk Preferences in the Small for a Large Population*

Hans-Martin von Gaudecker †
VU University Amsterdam and Netspar

Arthur van Soest
Tilburg University and Netspar

Erik Wengström
University of Copenhagen

April 18, 2008

Abstract

We analyse risk preferences using an experiment with small-stake gambles and real incentives in a representative sample of 1,422 Dutch respondents. Our structural econometric model incorporates four individual-level parameters: Utility curvature, loss aversion, preferences towards the timing of uncertainty resolution, and the propensity to choose randomly rather than on the basis of preferences. The model incorporates observed and unobserved heterogeneity. We find that socio-economic and demographic variables are significantly correlated with preference parameters but are of minor importance compared to the unobserved heterogeneity component. All four parameters contribute to explaining the choices of the individuals in our sample, but preferences towards the timing of uncertainty resolution play a smaller role than the others.

JEL Classification: C90, D81

Keywords: risk aversion, loss aversion, uncertainty resolution, Internet surveys

An appendix with additional material is available at:
`http://staff.feweb.vu.nl/hgaudecker/res/res_index.html`

*Financial support from the Dutch Science Foundation (NWO), the Swedish Institute for Banking Research (Bankforskningsinstitutet), and from the European Union under grant HPRN-CT-2002-00235 (RTN-AGE) is gratefully acknowledged. We thank the team of CentERdata, especially Marika Puumala, for their support with the experiments, as well as Morten Lau and Joachim Winter for very helpful remarks on the experimental design. The analysis benefitted from comments by James Banks, Syngjoo Choi, Erik Sørensen, Peter Wakker, and seminar participants in Mannheim, Amsterdam, Rotterdam, Tilburg, at the 8th RTN-AGE meeting in Paris, the XIIth FUR conference at LUISS in Rome, the 2006 ESA meetings in Nottingham and Tucson, and the 10th DNB Research Conference in Amsterdam.

†*Address for correspondence:* VU Department of Economics, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands. Email: hgaudecker@feweb.vu.nl.

1 Introduction

We describe and analyse an experiment on eliciting risk preferences using a representative sample from the Dutch population with more than 1,400 individuals. The main innovation of our work is that we estimate a structural model of the distribution of preferences in the population, distinguishing parameters for risk aversion, loss aversion, and a preference for the timing of uncertainty resolution. The rich nature of the dataset allows us to account for heterogeneity in all these parameters using a random coefficients model. In addition, we model heterogeneity in the tendency to make random optimization errors that explain why reported choices often are in conflict with fully rational decision making.

Analysing decisions under risk is an important theme in economic research. Economic theory has generated numerous models for it, starting with the early work of von Neumann and Morgenstern (1947). The class of models that combine linearity in probabilities with a non-linear utility function remains the workhorse in much of modern economics. Starting with Allais (1953), however, many violations of this basic model have been documented. Starmer (2000) reviews this literature. One particularly persistent feature in experiments is a greater sensitivity to losses than to gains of similar size. This feature was widely popularised through prospect theory (Kahneman and Tversky 1979) and defined formally as a kink of the utility function at the reference point by Köbberling and Wakker (2005). Heterogeneity in utility curvature has been documented in many studies (Hey and Orme (1994) provides an important example), but nearly all studies that incorporate loss aversion concentrate on mean or median parameters. Recent exceptions include Fehr and Götte (2005) and Gächter, Johnson, and Herrmann (2007) who assume linear utility on each side of the reference point. A novel aspect of our analysis is the joint estimation of individual-specific parameters measuring the utility curvature and the strength of the kink in a structural model.

A difference between many real life economic decisions and most economic experiments is a temporal separation between the decision and the resolution of uncertainty. For example, this is a characteristic of many asset allocation or insurance decisions. The timing of uncertainty resolution may matter for two reasons: A planning advantage of early resolution (Kreps and Porteus 1978) and anticipatory feelings such as hope and anxiety in case of late resolution (cf. Wu (1999) and Caplin and Leahy (2001)). Preferences for the timing of uncertainty resolution have received considerable empirical attention in the last decade, but the literature does not appear to have reached any definite conclusions yet (see Chew and Ho (1994), Ahlbrecht and Weber (1996), Lovallo and Kahneman (2000), Noussair and Wu (2006), Eliaz and Schotter (2007), or van Winden, Krawczyk, and Hopfensitz (2007)).

Other than standard experiments with immediate payoffs, all our incentives are paid out

three months after the experiment. The timing of uncertainty resolution is varied between the beginning and the end of this period. This allows us to identify preferences for the timing of uncertainty resolution and to estimate an adjusted version of the Kreps and Porteus (1978) model that explicitly distinguishes parameters for utility curvature, loss aversion, and preferences for the timing of uncertainty resolution.

Recently, there has been an increased interest in using economic experiments to draw inference on the distribution of economically important preference parameters in a broad heterogeneous population (Harrison, Lau, and Williams (2002), Bleichrodt, Pinto, and Wakker (2001), Andersen et al. (2008), among many others). There has been a growing concern that the standard recruitment procedure – an experimenter inviting college students via emails or posters – may restrict socio-demographic variation too severely as to allow for meaningful inference on the broad population of interest. Spurred by Harrison, Lau, and Williams (2002), this issue has been addressed in a number of recent field studies.

We conduct our experiment on a heterogeneous subject pool that is representative of the Dutch population. Moreover, we model heterogeneity in all preference parameters, as a function of observed as well as unobserved characteristics. Our experiment is based on the CentERpanel, an Internet-based household survey in the Netherlands. It has been used before for related questions: Donkers, Melenberg, and van Soest (2001) and Booij and van de Kuilen (2006) use hypothetical questions to estimate risk preference functionals. They correlate them with observed covariates, but do not consider idiosyncratic heterogeneity. Huck and Müller (2007) provide evidence that violations of Expected Utility Theory in the Allais paradox occur more frequently in some population subgroups than in others. We also briefly compare our experimental results with a parallel laboratory experiment with students from Tilburg University.

Most theories of utility are deterministic in nature. Provided knowledge of the relevant parameters, they imply a unique choice from any set of options except for very special cases of indifference. Wherever repeated or sufficiently similar choices allow a violation of revealed preference conditions, a fraction of subjects will violate them (see Hey and Orme (1994) and Choi et al. (2007) for two of the most prominent examples). Explanations for this include changing tastes, lack of motivation, and difficulties of understanding the problem. Empirical models have therefore added noise to the process that generates observed choices. There has been considerable progress on modeling this noise. Key contributions include Harless and Camerer (1994), Hey and Orme (1994), Loomes and Sugden (1995), Ballinger and Wilcox (1997), Loomes, Moffatt, and Sugden (2002), and the papers in a special issue of Experimental Economics (Starmer and Bardsley 2005).

We build on this literature and estimate individual-specific error propensities along with

the preference parameters. This allows us to assess how much confidence we can have in the estimates of individuals' preference parameters. Put differently, our estimates yield signal-to-noise ratios for any observed set of choices.

Our main results can be summarised as follows. Utility curvature and loss aversion turn out to be very important determinants of individuals' choices under risk. The influence of preferences towards the resolution of uncertainty is less important, but not negligible for everybody. While many people exhibit consistent choice patterns, some have very high error propensities. Errors are much more prevalent for the general population than for the students population represented in the laboratory. Parameters for utility curvature and loss aversion also vary with socio-demographics and this variation implies that students are not representative for the broader population. Still, the variation in preferences induced by our rich set of socio-demographic variables is small compared to the variance ascribed to the unobserved heterogeneity component. This implies that controls for individual preferences contain useful information in addition to socio-demographics and it would be desirable to make them more widely available for empirical work based upon micro-data from socio-economic surveys.

The remainder of this paper is organised as follows. In the next section, we describe the experimental design and motivate the subsequent analysis by describing the choice behaviour of five "benchmark" subjects, highlighting some common and some extreme examples of choice behaviour. In Section 3 we set up the theoretical model and describe the econometric strategy. Section 4 contains our results, describing average preference parameters and their socio-demographic correlates, the nature of unobserved heterogeneity, and the implications of the choices of the five selected subjects for their preference parameters. Section 5 concludes.

2 Data and Experimental Setup

We implemented the experiment in the CentERpanel, a Dutch household survey that is administered via the Internet. In order to avoid selection problems due to lack of Internet access, respondents without a computer are equipped with a set-top box for their television set (and with a TV if they do not have one). The panel consists of roughly 2,000 households who are representative of the Dutch population in terms of observable characteristics. It has rich background information on important demographic and socio-economic characteristics. Respondents are reimbursed regularly for their expenses connected with Internet use and we used the existing system for reimbursement to make payments to the panel members who participated in the experiment.

We provide a description of the covariates and the construction of the sample used for

estimation in Appendix A. We use data on 1,422 individuals in the CentERpanel. Parallel experiments were conducted at Tilburg University's economics laboratory with a total of 178 student participants. While these are not the focus of this paper, we provide some comparisons to results from these experiments.[1]

2.1 Experimental Design

Our experiment uses an adapted version of the well-established multiple price list format, applied earlier by, for example, Binswanger (1980), Tversky and Kahneman (1992), Holt and Laury (2002), and Harrison, Lau, and Rutström (2007). Andersen et al. (2006) provide a detailed description and we limit ourselves to a brief introduction. Each subject is shown four pairs of lotteries such as the ones presented in Figure 1. We call the components of these pairs option 'A' and option 'B', where option 'B' always involves as much or more risk than option 'A'. Subjects may opt for either option in each of the four choice tasks. The payoffs of both options do not change, but the probabilities of the high payoff in each option vary from 25 % to 100 % as one moves down the screen. The table is designed such that the expected value of option 'A' starts out higher but moves up slower than the expected value of option 'B'.

Utility-maximising participants switch at some point from option 'A' to option 'B' or choose option 'B' throughout the screen. If such consistent behaviour is observed, the subject is routed to a screen containing lotteries with the same payoffs, but a finer probability grid. Andersen et al. (2006) recommend using this method and call it "iterated multiple price list", other authors have used the term "chained method" for similar strategies (Wakker and Deneffe 1996). The grid now consists of steps of 10 percentage points located roughly between the subject's highest choice of 'A' and his lowest choice of 'B' on the first screen. For a set of lotteries that differ only in the probabilities we use the term "payoff configuration". All subjects were given the seven payoff configurations listed in Table 1. For each configuration, subjects made either four or eight decisions, depending on their answers on the first screen. Our data therefore constitute an unbalanced panel of 28 to 56 binary choices for each respondent.

A modification compared to previous studies is the inclusion of pie-charts describing the probabilities of the outcomes. Pilot experiments showed that this supplement to the verbal descriptions of the decision tasks was appreciated by subjects who were less familiar with probability judgements. Moreover, we restricted the number of decision tasks per screen from the usual 10 to 4 to avoid the need for scrolling. Subjects could not go back to revise

[1] See von Gaudecker, van Soest, and Wengström (2008) for an extensive analysis comparing the laboratory and the Internet experiments.

decisions. We tested the instructions thoroughly. Unlike in typical laboratory settings, there was no experimenter to answer questions. In order to compensate for this, subjects had access to the instructions and specially designed help screens throughout the experiment.

Participants were allocated to one of three different incentive treatments. One of these was entirely hypothetical, with lotteries based upon the payoff configurations in Table 1. The two treatments with real incentives involved a participation fee paid to everyone who completed the experiment. Additionally, for one in every ten participants in these two treatments, one lottery was randomly selected and played out, and the payoff of that lottery was paid out. .[2][3] In the high incentive treatment, the completion fee was 15 Euros, with lottery payoffs listed in Table 1. In the low incentive treatment, payoffs and participation fee were one third of those of the high incentive treatment. We allocated subjects to one of two randomly determined orderings of the seven sheets.

Some of the lotteries included negative outcomes and we use a zero lottery payoff as the natural reference point. We avoided negative overall payoffs by setting the participation fees equal to the maximum losses that could be incurred. It should be noted that this has the potential drawback that loss aversion might be underestimated, since some respondents may combine payoff and participation fee and use a different reference point so that they never can experience a loss. The fact that we still find strong evidence of loss aversion suggests that this is not so important. Also note that the less risky option 'A' always contained only weakly positive outcomes, so that subjects were able to avoid losses altogether.

All payoffs were made three months after the experiment. The timing of uncertainty resolution was set to either directly after the experiment or just before the payment was made. We emphasized to subjects that this concerned only the timing of uncertainty resolution and not the timing of payment: an entire screen of the introduction was devoted to this aspect and the timing of uncertainty resolution featured prominently at the top of each payoff configuration (see Figure 1).

[2] Andersen et al. (2007) find no significant difference between choices made in the case if one in ten participants is paid or if all participants are paid.

[3] The iteration introduces a distortion of incentive compatibility for some subjects. Sufficiently risk-tolerant participants (i.e. those whose preferences imply a switchpoint at or below 75%) had an incentive to switch at a higher probability for the high payoff because this would raise expected payoffs in the second step. This does not appear to be a major problem as the iteration procedure was not announced and we do not observe learning in the sense of increasing switch points over the course of the experiment

2.2 Heterogeneity in the Choice Data

In this section we highlight the features of our experimental data that motivate the economic modelling. We illustrate the heterogeneity in individual behaviour by describing the choices of five selected individuals, labelled respondents 1–5. Their choices are presented in in Figure 2. Some aggregate descriptive statistics can be found in von Gaudecker, van Soest, and Wengström (2008).

Panel A of Figure 2 presents the choices made by individual 1. The pattern is typical of the behaviour that we observe: While mostly consistent, the choices also exhibit some inconsistencies. This concerns the last three payoff configurations (5, 6, and 7), where we observe some nonmonotonicities (5 and 7) and different choices when faced with the same decision twice (6). The latter could also reflect indifference between the two options. There are more risky choices in payoff configurations 1 and 2 than in 3 and 4. Configurations 1 and 2 are characterised by early resolution and nonnegative payoffs. Behaviour on screens 1 and 2 still shows risk aversion – as in Holt and Laury (2002), risk neutrality would imply switches at 50% probability of the high payoff in all cases. A larger number of safe choices was made in payoff configurations 3 (with a probability of a negative payoff in option 'B' and early resolution in both options) and 4 (late resolution of option 'B', positive outcomes only).

The other four benchmark individuals have choice patterns that are less common though not exceptional – most observed choice patterns are somewhere in the range between those of individual 1 and one of the other individuals. The choices of individual 2 (Figure 2B) are very consistent and mostly safe – in all cases where the outcome set included a negative or zero payoff , individual 2 avoided this possibility by choosing the safer lottery. Individual 3 (Figure 2C) makes the same choices as individual 2 for payoff configurations 5 and 7, but the choices on the other screens suggest that these are generated from different underlying preferences: In the other payoff configurations the choices of individual 3 are much riskier than those of subject 2. This shows that a combination of moderate risk aversion, loss aversion, and preference for the early resolution of uncertainty leads to the strongly risk avoiding behaviour for payoff configurations 5 and 7.

Figure 2D shows the opposite to individual 2: A subject exhibiting risk-lovingness on all screens. There are enough such patterns in the data to require a specification of heterogeneous preferences that can accommodate this kind of behaviour. Finally, individual 5 in Figure 10A makes completely erratic choices without a clear pattern. Few individuals behave like this, but there are many intermediate cases with a pattern suggesting an error tendency between that of subjects 1 and 5. This wide range of choice patterns with inconsistencies makes careful

modelling of errors important.[4]

3 Theoretical Framework and Econometric Implementation

In this section, we first lay out the utility specifications that form the basis of our econometric analysis. We start with an expected utility of income specification and then incorporate loss aversion. In Section 3.2, we introduce a temporal component in a two-period model. While all payments are made in the second period, uncertainty may be resolved either in the first or the second period. We implement a parsimonious version of the Kreps and Porteus (1978) model in order to allow for preferences towards the timing of uncertainty resolution. The last step consists in developing an empirical model that allows for sufficient heterogeneity (Section 3.3).

3.1 A Simple Model of Choice Under Risk

We start from a standard expected utility formulation with an exponential utility function:

$$u(z,\gamma) = -\frac{1}{\gamma}e^{-\gamma z} \tag{1}$$

where $z \in \mathbb{R}$ denotes a lottery outcome and $\gamma \in \mathbb{R}$ is the Arrow-Pratt coefficient of absolute risk aversion. We prefer an exponential utility formulation over power utility because it lends itself better to the subsequent generalisations. In a separate appendix (available on line) we compare the results to alternative specifications using a power utility function.

The first extension of (1) is to incorporate loss aversion, following prospect theory (Kahneman and Tversky 1979) and in line with the widely recognised stylised fact that "losses loom larger than gains" (see, e.g., Starmer (2000) for a review). In line with the literature, we augment (1) with a loss aversion parameter $\lambda \in \mathbb{R}_+$:

$$u(z,\gamma,\lambda) = \begin{cases} -\frac{1}{\gamma}e^{-\gamma z} & \text{for } z \geq 0 \\ \frac{\lambda-1}{\gamma} - \frac{\lambda}{\gamma}e^{-\gamma z} & \text{for } z < 0 \end{cases} \tag{2}$$

The degree of loss aversion is measured by the ratio of the left and the right derivatives of the utility function at zero, as suggested by Köbberling and Wakker (2005). Prospect theory's original utility function is concave for gains and convex for losses. In contrast to this, (2) assumes the same type of curvature on the whole real line. This is motivated by the need to build (2) into a Kreps-Porteus environment (see below) and by some recent empirical results

[4] In our data, there is no clear cut-off point to exclude some individuals a priori from the sample, as is common in the literature (see, e.g., Abdellaoui, Barrios, and Wakker (2007) or Choi et al. (2007)).

that call prospect theory's original utility curvature findings for mixed gambles into question (Baltussen, Post, and van Vliet 2006). Ideally, we would estimate separate γ parameters for the gains and loss domains. However, our experiment does not have enough variation in negative outcomes to do this. In the Online Appendix, we present estimates based on alternative functional form assumptions that include prospect theory type preferences.

3.2 Preferences towards the Timing of Uncertainty Resolution

In order to model preferences towards the timing of uncertainty resolution, we adopt the general framework of Kreps and Porteus (1978). In line with our experimental setup we consider a two-period setting. All decisions are made in the first period and payments made in the second period. The outcome of a gamble is either revealed in period 1, directly after the choices have been made (early resolution), or at the time of the payments in period 2 (late resolution).[5]

Assume that agents first calculate period 2 utility for all outcomes based on a function $v(z, \cdot)$, where z is the payoff and the dot replaces preference parameters. Thereafter, they use a continuous and strictly increasing weighting function $h(\cdot)$ to calculate their first period utility, with period 2 utility as its argument. The period 1 utility of a degenerate lottery that gives a certain outcome in period 2 is then given by $h(v(z, \cdot))$. The evaluation of nondegenerate lotteries hinges on the timing of uncertainty resolution – the expectations operator is applied either to the weighted or unweighted period 2 utility. Formally, let V denote the period 1 utility function for lotteries with payoffs in period 1. The first period utility evaluation of a lottery π is then given by:

$$V(\pi) = \begin{cases} \mathbb{E}[h(v(z, \cdot))] & \text{for early uncertainty resolution} \\ h(\mathbb{E}[v(z, \cdot)]) & \text{for late uncertainty resolution} \end{cases} \tag{3}$$

Note that the expectations operator is always applied to the quantity that is known at the end of period one. If uncertainty resolves early, the decision-maker applies the weighting function to the utility of the specific outcomes of π. If the outcome of π remains uncertain until the second period, he applies the weighting function to its expected value. Kreps and Porteus (1978) show that h is strictly convex (strictly concave, linear) if and only if the decision maker always prefers early to late resolution (late to early, exhibits indifference). We

[5]Note that our motivation for modelling uncertainty resolution timing preferences is based on anticipatory utility. This concept does not seem to be captured well in the Kreps-Porteus model in settings where decisions take place in both periods because it imposes temporal consistency (Caplin and Leahy 2001). In our setup without decisions in period 2, however, it is general enough and provides an attractive way of incorporating static and dynamic lottery characteristics.

choose the following parsimonious "power function" specification of the weighting function:

$$h(v(z,\cdot)) \;=\; -S(-S\,v(z,\cdot))^{\rho^{-S}} \tag{4}$$

with $\rho \in \mathbb{R}_+$ and S denoting the sign operator given by:

$$S \;=\; \begin{cases} 1 & \text{for } \gamma \geq 0 \\ -1 & \text{for } \gamma < 0. \end{cases} \tag{5}$$

For $\rho > 1$, $h(\cdot)$ is convex and early resolution is preferred to late resolution. Indifference is obtained for $\rho = 1$, and late resolution is preferred for $\rho < 1$. We model the second period utility function as a slightly modified version of (2):

$$v(z,\gamma,\lambda,\rho) \;=\; \begin{cases} \max\{-\frac{\lambda}{\gamma},0\} - \frac{1}{\gamma}e^{-\gamma\rho^S z} & \text{for } z \geq 0 \\ \max\{-\frac{\lambda}{\gamma},0\} + \frac{\lambda-1}{\gamma} - \frac{\lambda}{\gamma}e^{-\gamma\rho^S z} & \text{for } z < 0 \end{cases} \tag{6}$$

The building blocks of (4) and (6) seem complicated because of the necessity to accommodate both types of utility curvature. The term $-\frac{\lambda}{\gamma}$ is added for risk lovers to assure that the weighting function $h(\cdot)$ can be applied, i.e. it guarantees that $v(z,\gamma,\lambda,\rho)$ is greater than zero for $\gamma < 0$. Including ρ^S in the exponent serves to retain the interpretation of $\gamma > 0$ as the coefficient of absolute risk aversion for early resolving lotteries on the positive domain. For such lotteries $V(\pi)$ collapses to $\mathbb{E}[u(\pi)]$ given in (2) if the subject is not risk loving. This implies that the distinction between risk aversion and uncertainty resolution timing preferences is identified for risk averse subjects if there are gambles on the positive domain and the timing of uncertainty resolution is varied.[6]

3.3 Econometric Implementation

Based on this specification of utility, we formulate structural econometric models that can be estimated by maximum likelihood. The models allow for individual heterogeneity in preference parameters and in the tendency to make errors. The heterogeneity can be captured by observed characteristics ("observed heterogeneity") or not ("unobserved heterogeneity"). Assume that individual $i \in \{1,\ldots,N\}$ faces $j \in \{1,\ldots,J_i\}$ dichotomous choices between two binary lotteries A and B, with $\pi^A_{ij} = \left(A^{low}_{ij}, A^{high}_{ij}, p^{high}_{ij}\right)$ denoting the low and high payoff and the probability of the high payoff for lottery A in choice situation j, and similarly $\pi^B_{ij} = \left(B^{low}_{ij}, B^{high}_{ij}, p^{high}_{ij}\right)$ for lottery B. Let $Y_{ij} = 1$ if the individual opts for π^B_{ij}

[6] This distinction is only approximately true for gambles with negative outcomes because of the additive term $\frac{\lambda-1}{\gamma}$ in (6). For risk lovers, the inclusion of $-\frac{\lambda}{\gamma}$ distorts the interpretation of γ by the same token. For the parameter values that we estimate, the magnitudes of the distortions are small. The Köbberling and Wakker (2005) definition of loss aversion continues to hold for period 2 utility.

(i.e., chooses lottery B in choice situation j) and $Y_{ij} = 0$ otherwise. Define the difference in certainty equivalents of the two lotteries in decision task j as:

$$\Delta\text{CE}_{ij} = \text{CE}(\pi_{ij}^B, \gamma_i, \lambda_i, \rho_i) - \text{CE}(\pi_{ij}^A, \gamma_i, \lambda_i, \rho_i),$$

where $\text{CE}(\pi_{ij}^k, \cdot)$, $k = A, B$ is the period one certainty equivalent of lottery π_{ij}^k given the utility function in (3) in combination with (4) and (6). It is easy to derive an exact analytical expression for CE under our functional form assumptions, see Appendix B.

A perfectly rational decision maker would choose option 'B' if and only if $\Delta\text{CE}_{ij} > 0$. As a first step to allow for stochastic decision making, we add error terms to the CE comparison and model the individual's choice as:

$$Y_{ij} = \mathbb{I}\left\{\Delta\text{CE}_{ij} + \tau\,\varepsilon_{ij} > 0\right\}, \tag{7}$$

where $\mathbb{I}\{\cdot\}$ denotes the indicator function. We assume that the ε_{ij} are independent of each other and of the random coefficients driving the utility function, and follow a standard logistic distribution. The parameter $\tau \in \mathbb{R}_+$ governs the individual's probability to make "mistakes"; the probability of such a mistake falls with the absolute value of ΔCE.

The use of certainty equivalents in (7) leads to a meaningful interpretation of ΔCE in monetary terms. This means that τ has an intuitive interpretation. For example, if the difference in valuations of two lotteries (i.e. the "cost" of making an error relative to the individual utility function parameters) is $\Delta CE = 10$ Euros and $\tau = 4$, the probability to choose the higher-valued lottery is .92. For $\Delta CE = 1$ Euro, it is only .56. Using certainty equivalents facilitates comparisons across subjects – using utility differences directly as, for example, Hey and Orme (1994), leads to additional dependence on the preference parameters and makes comparisons between subjects difficult.

In addition to adding the error terms $\tau\,\epsilon_{ij}$, we allow subjects choose at random in any given task, following Harless and Camerer (1994). The propensity to do so is governed by the individual specific parameter $\omega_i \in [0, 1]$ and the probability of the observed choice Y_{ij} of individual i in choice situation j, given all the individual specific parameters, is given by:

(8)
$$l_{ij}\left(\pi_{ij}^A, \pi_{ij}^B, Y_{ij}, \tau, \gamma_i, \lambda_i, \rho_i, \omega_i\right) = (1-\omega_i)\,\Lambda\left((2Y_{ij}-1)\frac{1}{\tau}\Delta\text{CE}_{ij}\left(\pi_{ij}^A, \pi_{ij}^B, \gamma_i, \lambda_i, \rho_i\right)\right) + \frac{\omega_i}{2},$$

where $\Lambda(t) = (1 + e^{-t})^{-1}$ stands for the cumulative standard logistic distribution

For the sake of a parsimonious and easily interpretable model, we restrict τ to be the same for all individuals while allowing subjects to vary in their probability to make random choices. Alternative error specifications are possible (see the references in Section 1) but beyond the

scope of the current paper. For example, one might argue that τ should also be individual specific, but in practice it will be difficult to estimate heterogeneity in τ and ω separately (although both are identified, in theory).

We us a random coefficients model in order to estimate the distribution of the individual-specific parameters γ_i, λ_i, ρ_i, and ω_i in the population. This is a natural way of incorporating observed and unobserved heterogeneity directly.[7] An alternative approach would be to first estimate the parameters for each individual separately (Hey and Orme 1994) and then regressing the results on socio-demographic characteristics (Dohmen et al. 2005). This approach does not work in our case since the decisions depend on several parameters and unbiased or consistent estimation of individual parameters is not possible. In contrast to the finite mixture model of Harrison and Rutström (2006), we use a continuous distribution of the parameters of interest. The reason is that finite mixture models have difficulties handling a large number of potential values for the parameters and a small set of values seems insufficient to explain the very heterogeneous choice behaviour illustrated in Section 2.2. Our modeling of unobserved heterogeneity is similar to that of Conte, Hey, and Moffatt (2008); a difference is thatwe include observed heterogeneity in addition.

In order to work with a concise notation, define

$$\eta_i = g_\eta(X_i^\eta \beta^\eta + \xi_i^\eta), \quad \eta_i \in \{\gamma_i, \lambda_i, \rho_i, \omega_i\} \tag{9}$$

where η_i denotes one of the four individual specific parameters, X_i^η are $1 \times K^\eta$ vectors of regressors, β^η are $K^\eta \times 1$ parameter vectors, and ξ_i^η are the unobserved heterogeneity components of the parameters. The first element of each X_i^η contains 1. The functions $g_\eta(\cdot)$ serve to impose the theoretical restrictions on the individual specific parameters. For γ, this is just the identity function; for λ and ρ, it is the exponential function, guaranteeing that these parameters are positive; and for ω, it is the logistic distribution function, guaranteeing that ω is always between 0 and 1. We write $g(X_i\beta + \xi_i)$ for the vector of these four functions.

For the high incentive treatments, we assume that $\xi_i = (\xi_i^\gamma, \xi_i^\lambda, \xi_i^\rho, \xi_i^\omega)'$ follows a jointly normal distribution independent of the regressors, or, in other words:

$$\begin{pmatrix} g_\gamma^{-1}(\gamma_i) \\ g_\lambda^{-1}(\lambda_i) \\ g_\rho^{-1}(\rho_i) \\ g_\omega^{-1}(\omega_i) \end{pmatrix} \sim N\left(\begin{pmatrix} X_i^\gamma \beta^\gamma \\ X_i^\lambda \beta^\lambda \\ X_i^\rho \beta^\rho \\ X_i^\omega \beta^\omega \end{pmatrix}, \Sigma'\Sigma \right), \tag{10}$$

where Σ denotes an upper triangular matrix of Cholesky-factors. The regressor matrix contains a dummy for the hypothetical high incentives treatment to capture the potential effect

[7] See Bellemare, Kröger, and van Soest (2008) for a similar modelling approach

of giving hypothetical versus real payoffs. Preliminary estimations showed that the difference between low and high incentive treatments is better captured by a multiplicative specification than by adding a low incentive dummy to X. For the low incentive treatment, we therefore multiply each parameter in (10) by a parameter $\beta^{\eta}_{\text{low incentive}}$. This implies that not only the means but also the variance of the transformed parameter distributions change if we switch from high to low incentive treatment.[8]

Defining $\xi^* = (\Sigma')^{-1}\xi$ we can express the individual likelihood contributions as:

$$l_i = \int_{\mathbb{R}^4} \left[\prod_{j=1}^{J_i} l_{ij}\left(\pi^A_{ij}, \pi^B_{ij}, Y_{ij}, \tau, g(X_i\beta + \xi^*)\right) \right] \phi(\xi^*) d\xi^* \tag{11}$$

where l_{ij} is the probability given in (9) and $\phi(\cdot)$ denotes the probability density function of $N(0, I)$. The log likelihood is given by the sum of the logs of l_i over all respondents in the sample and can be maximised by standard methods to obtain the maximum likelihood estimates. The integral in equation (11) does not have an analytical solution and we use standard simulation techniques to approximate it. In particular, we employ Halton sequences of length $R = 1000$ per individual (Train 2003). We employ the BFGS algorithm with numerical derivatives to maximise the likelihood function. The variance-covariance matrix of the parameter estimates is based on the outer product of gradients. Standard errors for transformed parameters are calculated using the delta method.

4 Results

We present our results in three stages. First, we focus on average parameters and the role of observed heterogeneity, i.e. we discuss the β^{η} estimates. Second, we describe the entire population distribution of the preference and error parameters. In Section 4.3, we come back to the choices by individuals 1-5 that we described in Section 2.2 and investigate what these choices tell us about their individual preference parameters and error propensities.

Because of the model's complexity it is not straightforward to present the results in an easily accessible manner. We opt for showing them in several ways. For the socio-demographic correlates, we use standard tables of the estimates. We plot the estimated marginal distributions of preference and error parameters. The magnitudes of the preference parameters are difficult to interpret directly. Following Choi et al. (2007), we terefore report the risk premia (i.e. $\text{RP}(\pi) = \mathbb{E}[\pi] - \text{CE}(\pi)$) of standardised gambles implied by certain parameter constellations. To remain in line with the range of our payoffs, we use the gambles $\pi^1 = (25, 65, .5)$ and

[8] The multiplicative specification was selected by first considering separate models for each incentive treatment. This issue would not arise if we were only interested in average parameters without modelling individual heterogeneity.

$\pi^2 = (-15, 25, .5)$, with both early and late resolution of uncertainty. Since $\mathrm{RP}(\pi) \in [-20, 20]$ for both lotteries, the risk premia are directly comparable and their difference illustrates the impact of the loss aversion coefficient.

4.1 Average Parameter Estimates and Demographic Correlates

The estimates for the four parameter vectors β^η, $\eta \in \{\gamma, \lambda, \rho, \omega\}$, are displayed in Tables 3-6. For the constant, they are transformed back to the original scale, i.e. $g_\eta(\beta_1^\eta)$. Its value represents the median preference or error parameter for the population subgroup defined by the left-out categories. The means and medians coincide for γ and means will be larger than medians for the remaining parameters. For the other X_k, $k > 1$ covariates, the tables show $g(\beta_1^\eta + \beta_k^\eta) - g(\beta_1^\eta)$. These are the partial effects of changing the covariate from 0 to 1 on the median of the conditional distribution of the parameter given benchmark values of the other covariates. Table 7 contains the estimates for Σ, τ, and the values of the likelihood function. Allowing for non-zero off-diagonal elements of Σ gave inaccurate estimates. In the paper we therefore only report the results imposing a diagonal structure on Σ. Substantive results were largely insensitive to allowing for a general variance-covariance matrix; see the Online Appendix, Tables 8–11.

Each table has four columns. The first column contains estimates from the CentERpanel experiment, including only an intercept and the incentive treatments as explanatory variables. In the next column we add the full set of covariates. The remaining two columns show the corresponding values for the laboratory sample. First consider the top left elements in Tables 3–6. They show that the median subject's utility function in the high incentive treatment is concave, as expected. As most of the literature, we find significant loss aversion. The parameter ρ is quite precisely estimated and not significantly different from 1, implying that the median respondent is almost indifferent between early and late uncertainty resolution. The magnitudes of γ and λ are substantial in terms of the implied risk premia: They amount to 5.97 Euros for π^1 and to 10.65 Euros for π^2. The impact of ρ is negligible.

The raw estimates are difficult to compare to those found in the literature because of different parametric forms used for the utility functions, but we can compare the risk premia for π^1. Holt and Laury's (2002) power-expo function estimates based on a wide range of payoffs imply a risk premium of about 2.79 Euros. The CRRA interval of their median subject's choices in their 20x treatment (with payoffs in the same range as ours) leads to an interval between 1.90 Euros and 3.20 Euros. Choi et al.'s (2007) estimates for the median subject imply risk premia of 3.34 Euros or 5.50 Euros for the Gul-CRRA and Gul-CARA cases, respectively. In a representative sample of the Danish population, Harrison, Lau, and Rutström's (2007) estimates on the basis of a CRRA functional lead to risk premia

of 3.15 Euros. Our own estimates of about six Euros are slightly higher than the Choi et al. (2007) estimates that use an exponential type utility function and substantially higher than the other estimates that are based on power utility type functionals (or, in the case of Holt and Laury (2002), more general functionals that are close to power utility given the particular parameter estimates). The higher risk premia for the exponential utility seem to be a general characteristic in these settings. We see the same drop in risk premia as in Choi et al. (2007) when we estimate a prospect theory type model based upon a power utility function (see the online appendix for details, Section E and Table 21). The risk premium for the median subject is then 3.28 Euros, closely in line with the findings in the other studies.

The median random choice propensity (ω_i) is about 8.3%. This has to be seen in relation to the estimate for τ shown in Table 7A, which is about 4.0. For the examples used above in Section 3.3, this means that if $\Delta CE = 10$ Euros the probability to choose the higher-valued lottery is .88, while if $\Delta CE = 1$ Euro, this probability is only 0.55. This corresponds to the substantial error rates found in other studies using non-student samples. See, for example, de Roos and Sarafidis (2006) who analyse game show data with very high incentives or Huck and Müller (2007). From an analysis comparing lab and Internet data, von Gaudecker, van Soest, and Wengström (2008) conclude that the high error rate is indeed due to the nature of the sample rather than the Internet environment.

The general picture is not very different for the hypothetical treatment: γ and ω are virtually the same as for the real incentives case and ρ is again very cose to 1. λ is larger than for the real treatment. A potential interpretation of this is the show-up fee paid in case of the real incentive treatment – in part, subjects may have taken this into account so that they do not see the negative lottery payoff as a loss. There is no change in the risk premium for π^1 if the uncertainty is resolved early. For π^2, it rises by about one Euro to 11.74 Euros.

The estimated curvature of the utility function is substantially larger in the low incentive treatment. This behaviour of a power utility coefficient when varying the payoff scale has been reported before by Holt and Laury (2002). In contrast to them, our goal is not to find a functional form that provides a good approximation to aggregate behaviour over varying stakes. We are mainly interested in heterogeneity across subjects at a given level of payoffs. As the comparisons to power utility formulations provided in the online appendix show, exponential utility is much better suited for this purpose in our particular case (section E.2). The loss aversion coefficient is lower in the low incentive than in the high incentive treatment, uncertainty resolution timing preferences are not significantly different. Finally, random choice probabilities are higher in the low incentive treatment, but the parameter τ is much smaller, so the implication for the total tendency to make a suboptimal choice is ambiguous.

The magnitudes of the estimated coefficients for the high incentive group are quite different

in the laboratory experiment than in the Internet experiment (cf. the third columns of Tables 3–6). The laboratory based estimates for the medians of γ and λ are substantially lower. We find no significant differences between Internet and lab estimates of uncertainty resolution timing preference. In line with the parameter estimates, estimated risk premia are substantially lower in the lab than in the Internet environment, with lab values reaching about 3.50 Euros for π^1 and 7.50 Euros for π^2. The median propensity to choose at random in the lab is very close to zero, in line with the relatively small errors that are typically observed for student samples (Hey and Orme 1994). We find similar differences between hypothetical and real incentives treatments in the lab as over the Internet, except for the loss aversion coefficient, where the difference vanishes in the lab experiment. The fact that we paid subjects a show up fee in the hypothetical laboratory treatment but not in the corresponding CentERpanel group might explain this – the participation fee is partly taken into account, also if the lotteries are purely hypothetical. The coefficients indicating differences between low and high incentive treatments for the lab experiment are largely in line with those for the Internet experiment. An exception is the large difference for ω: its estimate for the high incentive treatment in the laboratory is disproportionally small.

The coefficients for the socio-demographic correlates in the second and fourth columns do not necessarily have a causal interpretation. For example, education, income, wealth, and financial literacy measures may be driven by the same unobserved characteristics as risk preferences. We can also not distinguish between age and cohort effects. Gender differences may also have their causes in unobserved factors (see Croson and Gneezy (2004) for a survey of potential mechanisms). Hence the term "marginal effect" refers to a ceteris paribus difference and does not imply a causal effect of changing the covariate. We find that women are significantly more risk averse than men, both in the CentERpanel and the laboratory experiments. This finding is common in the literature, although not all of the evidence is clear-cut (Croson and Gneezy 2004). In the laboratory, women appear to be more loss averse and more likely to prefer early resolution of uncertainty than men. In the CentERpanel experiment, they are more likely to choose at random. Gender is the only characteristic by which we can stratify the laboratory subjects because of limited variation (age, education) or lack of data. Hence the remainder of this section is only about the CentERpanel experiment.

For risk aversion, we find a positive age and a negative education gradient. The associations with income and wealth do not reveal a clear pattern. Being the household's financial administrator is associated with lower risk aversion. Loss aversion peaks for ages between 35 and 44 and then declines. It also declines with household income. Finally, we find that the elderly are significantly less averse to late uncertainty resolution than younger age groups.

The findings concerning risk aversion are in line with those of existing studies, cf. (Donkers,

Melenberg, and van Soest 2001) or Dohmen et al. (2005); Harrison, Lau, and Rutström (2007) finds hardly any significant effects but this may be due to their moderate sample size; Benjamin, Brown, and Shapiro (2006) find a negative association between risk aversion and cognitive skills, which is consistent with our findings for education). The results for loss aversion contradict those of Gächter, Johnson, and Herrmann (2007), possibly due to differences in experimental design, utility specifications, or sample selection[9]

The most striking differences between socio-demographic groups are found for the error parameter. For low-educated persons older than 65 with low income and wealth, we estimate a random choice propensity of 40% at the median. A plausible explanation is differences in numeracy, see Banks and Oldfield (2007). Not surprisingly, significantly fewer errors are made by the young and highly educated subjects. For this group, the error rates come close to those estimated in the laboratory. Point estimates for income and financial literacy are insignificant, but error rates decrease with wealth. We also included dummies for the time used to complete the experiment. They have the expected effect: those who completed the experiment rapidly have higher error rates and those who take a lot of time make fewer errors.[10]

4.2 Distributions of the Preference and Error Parameters

Having established important differences between population subgroups in the parameters of interest we now turn to describing their entire distribution, taking the unobserved component of heterogeneity into account as well. In doing so, we largely rely on graphical representations and use the high incentive treatment as the benchmark.[11] The solid lines in Figure 3 depict the marginal population distributions of our model's four random parameters. The dashed lines are kernel density estimates of the distibutions of the conditional medians given demographic characteristics. It is evident that the demographic controls account only for a small part of the total variation in the preferences. For example, 90% of the medians for the risk preference parameter γ lie in a range that accounts for less than a third of the distribution of the parameter itself, including the unobserved heterogeneity component. Only 1% of the conditional medians of ω imply a random choice propensity larger than .34, compared to 21% of the values of *omega* themselves.

Figure 4 plots the risk premia for $\pi^1 = (25, 65, .5)$ and $\pi^2 = (-15, 25, .5)$ implied at

[9]Gächter, Johnson, and Herrmann (2007) use one-shot real gambles to calculate bounds on loss aversion, assuming a piece-wise linear utility function and a sample of customers of a (mostly luxury) car manufacturer.

[10]Interestingly, we find no such effect in the laboratory sample.

[11]Since the allocation to treatment was random, the underlying preference parameter distributions will be the same in all three treatment groups. Differences in the estimated distributions that arise from the experimental design should therefore not be present in interpersonal comparisons.

various quantiles of the parameter distributions, showing that the variation in preference parameters implies substantial heterogeneity in terms of risk premia. The horizontal lines in Panels A-B plot the risk premia for the median preference parameters and for early resolution of uncertainty. The two sets of bars show the impact of moving each parameter separately to its 10th or 90th percentile.[12] For π^1, the median subject demands a risk premium of about six Euros (see Panel A). Changing γ to its 10th percentile gives a risk premium of -3.30 Euros, while at the 90th percentile it is 11.90 Euros. The risk premium of π_1 does not depend on λ or ρ. The picture is quite different for RP(π^2), as shown in Panel B. The baseline risk premium is now 10.65 Euros and the largest heterogeneity stems from the loss aversion parameter. The risk premium becomes negative when λ is set to its 10th percentile. Because of the high baseline value of λ, the effect of increasing it is much less. Finally, we see that due to the additive term in the definition of $v(z, \gamma, \lambda, \rho)$ in (6) for $z < 0$, there are slight changes in the risk premium when ρ is changed, in spite of the fact that only early resolution of uncertainty is considered.

Panels C and D depict the case of late resolution. The median subject is almost indifferent with respect to the uncertainty resolution timing, risk premia are about 13 Cents lower than for early resolution. The only important difference with respect to the early resolution case concerns the impact of changing ρ: Moving it to the 10th or 90th percentile now has an impact of up to 3,70 Euros on the risk premia. This effect is still small, however, compared to the heterogeneity due to the other parameters.

We conclude from this, that there is a lot of variation that is missed if the focus is on median parameters. This statement is hardly changed when one controls for socio-demographic covariates – i.e. having estimates of group medians (or means) rather than population medians hardly reduces the importance of the unobserved heterogeneity component. Put differently, the individual choices give us much more information than what is captured by socio-demographic groups. This will become even clearer in the next section. In line with this, accounting for the demographics hardly changes the estimated population distributions of the parameters. Those based on columns 1 and 2 of Tables 3-6 are virtually identical (see also the online appendix, (Figure 11).

Figure 5 illustrates the differences in the parameter distributions between the CentERpanel and the laboratory samples. The most salient difference is the much larger heterogeneity in the general population for all parameters except ρ. For the utility curvature parameter, this mirrors findings by Andersen et al. (2007); 95% of the laboratory subjects are estimated to have $\gamma \in [-.018, .049]$, whereas only slightly more than two thirds of CentERpanel subjects have a γ in this range. For the loss aversion parameter, only 1% of subjects in the student

[12] Table 21 in the online appendix contains the corresponding exact figures.

sample have a $\lambda > 10$, compared to more than 20 % in the CentERpanel sample. Finally, while 88% of the participants in the laboratory have a random choice probability below 10%, this fraction is only slightly more than one half in the CentERpanel.

4.3 Choices and Preferences at the Individual Level

The results of preceding section have revealed large differences in risk-taking behaviour across subjects of which only a small share is explained by differences in the subjects' socio-demographic characteristics. In this section we show how much information the choices in the experiment provide on the subjects'preference parameters. We do this for the five benchmark individuals described in Section 2.2. Our approach is similar to that of Revelt and Train (2000). The ("posterior") distribution of preference parameters of each respondent is derived conditional on the individual characteristics and the observed choices of that respondent and the estimated ("prior") distribution of the preference parameters.[13] For convenience, we reproduce the graphs with the raw choices made by the five individuals described in Section 2.2 in Panels A of Figures 6-10. Panels B-D contain plots of both the unconditional (prior) distribution $F\left(\eta_i \mid X_i, \hat{\beta}, \hat{\Sigma}, g_\eta\right)$ and the conditional (posterior) distribution $F\left(\eta_i \mid Y_i, X_i, \hat{\beta}, \hat{\Sigma}, g_\eta\right)$. The smaller the area under a graph, the more precise is the (posterior) estimate of the parameter. The general picture that emerges is that the γ_i can be estimated very accurately. For the other parameters, the quantitative implications for the parameters themselves are often imprecise, but we can still draw meaningful conclusions in terms of risk premia.

Consider Figure 6 with the posteriors for individual 1 in Section 2.2. Her risk aversion parameter is very likely near the mean of the unconditional distribution – the 10% and 90% quantiles of the marginal posterior distribution imply risk premia for π^1 of 4.61 Euros and 7.67 Euros, respectively (in line with the analysis above, all other values are set to their posterior median. Tables 22–25 contain the risk premia for the four individuals at various quantiles of their distributions). The same quantiles for the prior distribution (given the individual characteristics, imply risk premia of -3.14 Euros and 11.59 Euros. Hence knowing this subject's choices helps enormously to determine her preferences. For the loss aversion parameter, the distance between the first and ninth deciles shrinks from more than 14 Euros in the prior to less than six Euros when conditioning on the subject's choices. Panel D shows that there is less than a 10% chance that individual 1 prefers early to late resolution while the corresponding prior probability for her socio-economic group is more than 40%. Finally, Panel E reveals that her random choice propensity is in the medium range of the population distribution. For all four parameters, conditioning on the choices therefore makes the implied

[13] This is still conditional on the individual characteristics included in the vector of covariates X_i.

parameter ranges much tighter than conditioning on covariates only.

As expected from her choices, individual 2's risk aversion parameter is in the upper range of the distribution. The risk premia for π^1 implied by the parameters at the posterior first and ninth deciles are 9.37 Euros and 12.54 Euros, respectively. The chance that her loss aversion parameter is less than the group median of 4.4 lies slightly above 10%. The 90% quantile is 75 which makes sense since in all gambles with nonzero probability of a negative outcome this individual chose the safe option. The effect of this on the risk premia is surprisingly small, those for π^2 vary between 13.73 Euros and 14.47 Euros as λ moves from 4.3 to 75. The conditional distribution of ρ almost tracks the population distribution, implying that individual 2's choices provide very little information on her value of ρ. Panel E shows that because her choices are highly consistent, her error propensity is probably very low: With 95% probability, it is below 7%.

Individual 3 (Figure 8) made moderately risk averse choices if payoffs were nonnegative and resolution of uncertainty was early. For late resolution and potentially negative payouts, she always avoided the risky outcome. This choice pattern is reflected in the posterior distributions: Her utility curvature parameter is somewhat below the average with the first and ninth decile at .019 and .031 and corresponding risk premia for π^1 of 3.74 Euros and 5.87 Euros. The loss aversion parameter is more precisely estimated than for the previous individual since some of person 3's choices imply an upper bound on λ. The first and ninth deciles are at 6.59 and 28.22. The posterior median of ρ is estimated to be at 1.8 with the first and ninth decile at 1.34 and 2.44. This means that the risk premium of 4.77 Euros for π^1 in the early resolution case rises to between 6.40 Euros and 9.95 Euros if uncertainty resolves late.

Figure 9B clearly shows that the choices of individual 4 imply that he is risk-loving. He is very likely to be neither loss averse nor does he prefer early resolution of uncertainty. He has a low tendency to make suboptimal choices. The posterior distributions of this individual's parameters are quite narrow and directly reflect his choices. On the other extreme, individual 5's choices are hardly informative on her preferences (Figure 10). The marginal distributions conditioning on her choices are just as dispersed as the unconditional distributions. All we can say is that her behaviour is probably driven by a tendency to choose randomly instead of on the basis of the economic model (panel E).

5 Conclusions

We have described a large scale experiment on decision making under risk using a representative sample of a broad population. We have analyzed the experimental data using a structural

empirical model, disentangling preference parameters for utility curvature, loss aversion, and preference for the timing of uncertainty resolution, and allowing for several types of errors.

Our model requires a number of specification choices. To save space, we have only presented the results of our favourite model. In the Online Appendix (Section E) we further motivate our specification choices, showing that our model performs better than several alternatives. We also show that our substantive results are generally robust to the specification choice that we have made. For example, not accounting for early or late resolution preferences (setting ρ equal to 1) or changing the assumptions on the utility curvature to prospect theory type prefernces does not affect the estimates of the average risk premia for our benchmark lotteries very much. Using power utility rather than an exponential utility function leads to a drop similar to the one observed by Choi et al. (2007).

Our main finding is that risk preferences in the population are very heterogeneous, and only a small part of this heterogeneity can be captured with standard covariates such as age, gender and education level. Our structural model combined with the rich data with a large number of respondents making repeated choices appears to be a useful tool to handle the heterogeneity. Our four main parameters of interest (the three preference parameters and the tendency to choose purely at random instead of on the basis of utility maximisation) are modelled as random coefficients, and we find substantial dispersion in all of them. For example, even though we find that the timing of uncertainty resolution does not matter much for the median respondent, our estimates imply that there are groups in the population that clearly prefer early resolution and other groups that prefer late resolution. This is in line with the mixed evidence for temporal lotteries that has been accumulated so far in the literature.

Our structural model is particularly appropriate to analyse the informational content of each subject's choices for the parameters of that subject. This is shown by comparing the posterior distribution of some subjects given their choices to the prior distribution (given covariates only). We find that the choices are generally very informative about individual preference parameters, except, as expected, in cases where the choice data show that the subjects' choices are probably completely random.

Another lesson concerns the design of economic experiments. We have shown that the observed choices of many subjects were not always informative on their preferences due to the optimisation errors, particularly in the Internet experiment but also amongst students in the laboratory. One-shot questions aimed at "eliciting" individual preferences such as asking for the certainty equivalent of a lottery may lead to false precision in the estimates: Mistakes cannot be identified and random answers are taken for true preferences. In our opinion, an experimental design should account for inconsistent behaviour, for example through repeated choice tasks that generate overidentification. Choi et al. (2007) is a very elegant example

of such an approach. We have shown that given repeated choice data and our structural framework, it is fairly simple to construct a measure of confidence in an individual-level preference parameter. This may not be the most satisfactory answer, but it is certainly better than an estimate that seems precise but is not.

Our experiment works with relatively small stakes and the results do not necessarily apply to decisions involving much larger stakes often made in real life. For example, we find that the impact of the timing of uncertainty resolution is small for the large majority of the population. This finding may be confined to settings with comparatively small incentives. If incentives are increased sufficiently, in particular such that the planning advantage of early resolution becomes relevant, the timing of uncertainty resolution may also matter more. On the other hand, given the typically small incentives in most economic experiments, we take our result as good news for the standard practice of working with atemporal settings.

As demonstrated by Rabin (2000), the degree of utility curvature that is observed in our experiment cannot be expected to be directly useful for life-cycle decisions that involve much larger stakes. It remains an empirical question whether the same behavioural processes drive decisions in the small and in the large. This mirrors a debate in survey research methodology on the best measure of idiosyncratic preferences. The most common ones include large-scale hypothetical consumption choices (Barsky et al. 1997), self-rated risk tolerance in general and across different domains, or hypothetical investment choices (Dohmen et al. (2005) contains a combination of the latter). Our experiment adds another candidate to this list and a population representative Internet panel such as the CentERpanel provides an environment that is rich enough to compare different measures in terms of their predictive power for economic decisions of interest. The large variation in individual preferences that we have found certainly suggests that some of the heterogeneity in economic choices can be explained from heterogeneity in preferences. Using individual specific parameter predictions based upon experimental choices and a structural model such as ours seems a promising avenue to analyse this.

References

Abdellaoui, M., C. Barrios, and P. P. Wakker (2007): "Reconciling Introspective Utility with Revealed Preference: Experimental Arguments Based on Prospect Theory," *Journal of Econometrics*, 138(1), 356–378.

Ahlbrecht, M., and M. Weber (1996): "The Resolution of Uncertainty: An Experimental Study," *Journal of Institutional and Theoretical Economics*, 152, 593–607.

Allais, M. (1953): "Le Comportement de l'Homme Rationelle Devant le Risque: Critique des Postulats et Axiomes de L'École Américaine," *Econometrica*, 21, 593–607.

Andersen, S., G. W. Harrison, M. I. Lau, and E. E. Rutström (2006): "Elicitation Using Multiple Price List Formats," *Experimental Economics*, 9(4), 383–405.

——— (2007): "Preference Heterogeneity in Experiments: Comparing the Field and Laboratory," Centre for Economic and Business Research Discussion Paper No. 2005-03, Copenhagen.

——— (2008): "Eliciting Risk and Time Preferences," *Econometrica*, forthcoming.

Ballinger, T., and N. Wilcox (1997): "Decisions, Error and Heterogeneity," *The Economic Journal*, 107(443), 1090–1105.

Baltussen, G., T. Post, and P. van Vliet (2006): "Violations of Cumulative Prospect Theory in Mixed Gambles with Moderate Probabilities," *Management Science*, 52, 1288–1290.

Banks, J., and Z. Oldfield (2007): "Understanding Pensions: Cognitive Function, Numerical Ability and Retirement Saving," *Fiscal Studies*, 28(2), 143–170.

Barsky, R. B., F. T. Juster, M. S. Kimball, and M. D. Shapiro (1997): "Preference Parameters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Study," *Quarterly Journal of Economics*, 112(2), 537–579.

Bellemare, C., S. Kröger, and A. van Soest (2008): "Measuring Inequity Aversion in a Heterogeneous Population Using Experimental Decisions and Subjective Probabilities," *Econometrica*, forthcoming.

Benjamin, D., S. A. Brown, and J. Shapiro (2006): "Who is "Behavioral"? Cognitive Ability and Anomalous Preferences," Mimeo, Harvard University.

Binswanger, H. P. (1980): "Attitudes Towards Risk: An Experimental Measurement in Rural India," *American Journal of Agricultural Economics*, 62, 395–407.

Bleichrodt, H., J. L. Pinto, and P. P. Wakker (2001): "Making Descriptive Use of Prospect Theory to Improve the Prescriptive Use of Expected Utility," *Management Science*, 47(11), 1498–1514.

Booij, A. S., and G. van de Kuilen (2006): "A Parameter-Free Analysis of the Utility of Money for the General Population under Prospect Theory," University of Amsterdam, CREED Working Paper.

Caplin, A., and J. Leahy (2001): "Psychological Expected Utility Theory and Anticipatory Feelings," *Quarterly Journal of Economics*, 116, 55–79.

Chew, S. H., and J. L. Ho (1994): "Hope: An Empirical Study of Attitude Toward the Timing of Uncertainty Resolution," *Journal of Risk and Uncertainty*, 8, 267–288.

Choi, S., R. Fisman, D. Gale, and S. Kariv (2007): "Consistency and Heterogeneity of Individual Behavior under Uncertainty," *American Economic Review*, 97(5), 1921–1938.

Conte, A., J. D. Hey, and P. G. Moffatt (2008): "Mixture Models of Choice Under Risk," *Journal of Econometrics*, forthcoming.

Croson, R., and U. Gneezy (2004): "Gender Differences in Preferences," Mimeo, The Wharton School of the Universtiy of Pennsylvania.

de Roos, N., and Y. Sarafidis (2006): "Decision Making Under Risk in Deal or No Deal," Mimeo, University of Sydney. Available at SSRN: http://ssrn.com/abstract=881129.

Dohmen, T., A. Falk, D. Huffman, U. Sunde, J. Schupp, and G. G. Wagner (2005): "Individual Risk Attitudes: New Evidence from a Large, Representative, Experimentally-Validated Survey," IZA Discussion Paper 1730.

Donkers, B., B. Melenberg, and A. van Soest (2001): "Estimating Risk Attitudes Using Lotteries; A Large Sample Approach," *Journal of Risk and Uncertainty*, 22(2), 165–195.

Eliaz, K., and A. Schotter (2007): "Experimental Testing of Intrinsic Preferences for NonInstrumental Information," *American Economic Association - Papers and Proceedings*, 97(2), 166–169.

Fehr, E., and L. Götte (2005): "Do Workers Work More If Wages Are High? Evidence from a Randomized Field Experiment," University of Zurich, Institute for Empirical Research in Economics Working Paper No. 125.

Gächter, S., E. J. Johnson, and A. Herrmann (2007): "Individual-Level Loss Aversion in Riskless and Risky Choices," IZA Discussion Paper 2961.

Harless, D. W., and C. F. Camerer (1994): "The Predictive Utility of Generalized Expected Utility Theories," *Econometrica*, 62(6), 1251–1289.

Harrison, G., and E. Rutström (2006): "Expected Utility Theory and Prospect Theory: One Wedding and A Decent Funeral," University of Central Florida, Discussion Paper 05-18.

Harrison, G. W., M. I. Lau, and E. E. Rutström (2007): "Estimating Risk Attitudes in Denmark: A Field Experiment," *Scandinavian Journal of Economics*, 109(2), 341–368.

Harrison, G. W., M. I. Lau, and M. B. Williams (2002): "Estimating Discount Rates in Denmark: A Field Experiment," *American Economic Review*, 92(5), 1606–1617.

Hey, J. D., and C. Orme (1994): "Investigating Generalizations of Expected Utility Theory Using Experimental Data," *Econometrica*, 62(6), 1291–1326.

Holt, C. A., and S. K. Laury (2002): "Risk Aversion and Incentive Effects," *American Economic Review*, 92, 1644–1655.

Huck, S., and W. Müller (2007): "Allais for All: Revisiting the Paradox," Mimeo, University of Tilburg. Available at: http://center.uvt.nl/staff/muller/Allais071009.pdf.

Kahneman, D. V., and A. V. Tversky (1979): "Prospect Theory: An Analysis of Decision Under Risk," *Econometrica*, 47, 263–291.

Köbberling, V., and P. P. Wakker (2005): "An Index of Loss Aversion," *Journal of Economic Theory*, 122, 119–131.

Kreps, D. M., and E. L. Porteus (1978): "Temporal Resolution of Uncertainty and Dynamic Choice Theory," *Econometrica*, 46, 185–200.

Loomes, G., P. G. Moffatt, and R. Sugden (2002): "A Microeconometric Test of Alternative Stochastic Theories of Risky Choice," *Journal of Risk and Uncertainty*, 24(2), 103–130.

Loomes, G., and R. Sugden (1995): "Incorporating a Stochastic Element Into Decision Theories," *European Economic Review*, 39, 641–648.

Lovallo, D., and D. Kahneman (2000): "Living with Uncertainty: Attractiveness and Resolution Timing," *Journal of Behavioral Decision Making*, 13(2), 179–190.

Noussair, C., and P. Wu (2006): "Risk Tolerance in the Present and the Future: An Experimental Study," *Managerial and Decision Economics*, 27(6), 401 – 412.

Rabin, M. (2000): "Risk Aversion and Expected Utility Theory: A Calibration Theorem," *Econometrica*, 68(5), 1281–1292.

Revelt, D., and K. Train (2000): "Customer-Specific Taste Parameters and Mixed Logit: Households' Choice of Electricity Supplier," University of California at Berkeley, Economics Working Paper E00-274.

Starmer, C. (2000): "Developments in Non-Expected Utility Theory: The Hunt for a Descriptive Theory of Choice under Risk," *Journal of Economic Literature*, 38(2), 332–382.

Starmer, C., and N. Bardsley (2005): "Introduction to the Special Issue: Exploring the Error in Experimental Economics," *Experimental Economics*, 8(4), 295–299.

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

Tversky, A. V., and D. V. Kahneman (1992): "Advances in Prospect Theory: Cumulative Representation of Uncertainty," *Journal of Risk and Uncertainty*, 5(4), 297–323.

van Winden, F., M. Krawczyk, and A. Hopfensitz (2007): "Investment, Resolution of Risk, and the Role of Affect," Paper presented at the 10th DNB Research Conference, Amsterdam.

von Gaudecker, H.-M., A. van Soest, and E. Wengström (2008): "Selection and Mode Effects in Risk Preference Elicitation Experiments," IZA Discussion Paper 3321.

von Neumann, J., and O. Morgenstern (1947): *The Theory of Games and Economic Behavior*. Princeton University Press, Princeton, 2nd edn.

Wakker, P., and D. Deneffe (1996): "Eliciting von Neumann-Morgenstern Utilities When Probabilities are Distorted or Unknown," *Management Science*, 42(8), 1131–1150.

Wu, G. (1999): "Anxiety and Decision Making with Delayed Resolution of Uncertainty," *Theory and Decision*, 46, 159–199.

A Data Description

On most weekends throughout the year, CentERpanel respondents are asked to answer questions. Among other things, the Dutch National Bank Household Survey (DHS) is administered via the CentERpanel. Its questionnaire resembles standard household surveys and puts special emphasis on financial variables. We link our experimental data to the survey answers and thus have access to a rich amount of background variables. We now describe the variables employed in our analysis. Descriptive statistics can be found in Table 2. The selected set of variables is by no means an exhaustive list of those available in the CentERpanel, but other socio-demographic variables such as labour market status or family characteristics neither exhibited significant correlations with observed behaviour nor did they alter the results on the included variables substantially.

As the most important demographic characteristics, we use gender and age. We construct dummy variables for ages 35 to 44, ages 45 to 54, ages 55 to 64, and age 65 and older. Individuals aged 18 to 34 constitute the left-out category. Furthermore we make use of educational attainment in four categories: primary and lower secondary education (left-out category), higher secondary education and intermediate vocational training, higher vocational training, and university education. In the last category we also include university students.

We use two measures of households' financial status: Income and wealth. We construct a measure of total net annual household income and break it up into three categories: Below $22,000$ Euros, between $22,000$ Euros and $40,000$ Euros, and more than $40,000$ Euros. We construct total wealth as the sum of asset holdings in all categories recorded in the DHS (savings or deposit accounts, bonds, stocks, employer-sponsored savings plans, funds, owner-occupied and other housing property, subtracting debts and mortgages). We construct four dummy variables from this measure, please refer to Table 2 for the thresholds. We further use two proxies for financial literacy: Whether the individual is in charge of financial matters in the household; and whether the subject rates him- or herself as financially knowledgeable or very financially knowledgeable on a four-point scale (the other categories are more or less knowledgeable and not knowledgeable).

The final set of variables relates directly to the experiment. First, the incentive treatment the subject was randomly allocated to. Second, we divide the observed completion times of the experiment into three categories.

We conducted the CentERpanel experiment in November and December of 2005. In total, 2,299 persons logged into the system. Some 291 respondents chose not to participate after the introductory screen which contained an explicit non-participation option, another 80 subjects dropped out along the way. Finally, we excluded those 138 persons who went through the

whole experiment in less than 5:20 minutes, which was the minimum duration observed in the laboratory. See von Gaudecker, van Soest, and Wengström (2008) for more details and an investigation of selection issues. This leaves us with a sample of 1,790 subjects in the experiment. However, not all of these took part in the DHS earlier in 2005. In particular, educational attainment is missing for two subjects and the financial module (wealth and the financial literacy variables) is not available for 366 subjects. We estimated specifications that did not include these financial variables with the restricted and the unrestricted sample, results did not change much. Hence we stick to the sample with 1,422 subjects on whom we have complete information available throughout the analysis. These respondents made a total of 73,084 binary decisions.

Parallel experiments were conducted at Tilburg University's economics laboratory in September 2005 and May 2006. Payments and procedures were similar to the CentERpanel experiment, except for a show-up fee worth 5 Euros paid for recruitment reasons in the hypothetical treatment. In a further treatment specific to the laboratory, we varied the presence of an experimenter. Since we could not detect any differences between these treatments (von Gaudecker, van Soest, and Wengström 2008), we do not touch upon the issue here.

B Certainty Equivalents

This appendix shows the formulas for the certainty equivalent (CE) of a temporal lottery π in terms of period one utility as laid out in Section 3 is given by:

$$h(v(\mathrm{CE})) \quad = \quad V(\pi).$$

Solving this for the certainty equivalent leads to:

$$\mathrm{CE} \quad = \quad v^{-1}\left(h^{-1}\left[V(\pi),\rho\right],\gamma,\lambda,\rho\right).$$

In our particular framework, we have the following:

$$\begin{aligned} h^{-1}(y,\cdot) &= -S(-S\,y)^{\rho^S} \\ v^{-1}(y,\cdot) &= \begin{cases} -\frac{\ln(\gamma\max\{-\frac{\lambda}{\gamma},0\}-\gamma y)}{\gamma\rho^S} & \text{for } y \geq \max\{-\frac{\lambda}{\gamma},0\}-1/\gamma \\ -\frac{\ln(\gamma\max\{-\frac{\lambda}{\gamma},0\}+\lambda-1-\gamma y)-\ln(\lambda)}{\gamma\rho^S} & \text{for } y < \max\{-\frac{\lambda}{\gamma},0\}-1/\gamma \end{cases} \end{aligned}$$

The certainty equivalent of a gamble π is hence given by:

$$\mathrm{CE}(\pi) \quad = \quad \begin{cases} -\dfrac{\ln\left(\gamma \max\{-\frac{\lambda}{\gamma},0\}+\gamma\, S\,(-S\,V(\pi))^{\rho^S}\right)}{\gamma\rho^S} & \text{for } V(\pi) \geq \max\{-\frac{\lambda}{\gamma},0\} - 1/\gamma \\ -\dfrac{\ln\left(\gamma \max\{-\frac{\lambda}{\gamma},0\}+\lambda-1+\gamma\, S\,(-S\,V(\pi))^{\rho^S}\right)-\ln(\lambda)}{\gamma\rho^S} & \text{for } V(\pi) < \max\{-\frac{\lambda}{\gamma},0\} - 1/\gamma \end{cases}$$

C Tables

Table 1: Characteristics of the Seven Payoff Configurations

Payoff Configuration	**Uncertainty Resolution, A**	**Payoff Low, A**	**Payoff High, A**	**Uncertainty Resolution, B**	**Payoff Low, B**	**Payoff High, B**
1	early	27	33	early	0	69
2	early	39	48	early	9	87
3	early	12	15	early	-15	48
4	early	33	36	late	6	69
5	early	18	21	late	-9	54
6	early	24	27	early	-3	60
7	late	15	18	late	-12	51

Note: These values were shown in the high incentive and hypothetical treatments. For the low incentive treatment they were divided by three. The order was randomised.

Table 2: Descriptive Statistics on the Covariates

Variable Description	**Fraction**
Female	0.45
Age 16-34 years	0.24
Age 35-44 years	0.18
Age 45-54 years	0.23
Age 55-64 years	0.19
Age 65 years and older	0.17
Primary / lower secondary education	0.31
Higher secondary education / Intermediate Vocational Training	0.32
Higher Vocational Training	0.25
University Degree / University Student	0.11
Total net annual household income below 22k Euros	0.33
Total net annual household income $\in$ [22k Euros; 40k Euros)	0.49
Total net annual household income at least 40k Euros	0.18
Total wealth below 10k Euros	0.32
Total wealth $\in$ [10k Euros; 50k Euros)	0.14
Total wealth $\in$ [50k Euros; 200k Euros)	0.31
Total wealth at least 200k Euros	0.23
Respondent is the household's financial administrator	0.65
Respondent rates himself/herself as financially knowledgeable	0.26
High incentive treatment	0.31
Hypothetical treatment	0.32
Low incentive treatment	0.37
Took less than 9 minutes to complete the experiment	0.20
Took between 9 and 18 minutes to complete the experiment	0.55
Took more than 18 minutes to complete the experiment	0.24

Note: Number of Observations is 1,422. All variables are dichotomous variables, fractions may not sum to one because of rounding errors.

Table 3: Estimated Risk Attitudes (γ)

	CentERpanel		Laboratory	
Covariate	**(4),(6)**	**(4),(6)**	**(4),(6)**	**(4),(6)**
Constant	0.0322***	0.0317***	0.0181***	0.0126***
	(0.0010)	(0.0028)	(0.0022)	(0.0027)
Female		0.0080***		0.0064*
		(0.0015)		(0.0034)
Age 35-44		-0.0004		
		(0.0022)		
Age 45-54		0.0046**		
		(0.0023)		
Age 55-64		0.0035		
		(0.0026)		
Age 65+		0.0150***		
		(0.0026)		
Hi Sec Educ / Int Voc Train		-0.0064***		
		(0.0018)		
Higher Voc Train		-0.0020		
		(0.0021)		
University		-0.0164***		
		(0.0026)		
Income EUR 22k-40k		-0.0013		
		(0.0016)		
Income EUR 40k+		0.0006		
		(0.0023)		
Wealth EUR 10k-50k		0.0088***		
		(0.0022)		
Wealth EUR 51k-200k		-0.0028		
		(0.0020)		
Wealth EUR 201k+		-0.0033		
		(0.0024)		
HH Financial Admin		-0.0034**		
		(0.0017)		
Financially Knowledgeable		-0.0003		
		(0.0018)		
Hypothetical Treatment	-0.0004	0.0008	0.0013	0.0031
	(0.0015)	(0.0016)	(0.0033)	(0.0036)
Low Incentive Treatment †	2.78***	2.77***	2.84***	2.84***
	(0.0869)	(0.0873)	(0.275)	(0.253)

Note: Number of Observations is 1,422 (CentERpanel) and 178 (Laboratory), respectively. Estimation follows (11) based on the utility functions in the column headers. Regression coefficients are transformed back to the original scale by inverting the function on the left-hand side of (10). Except for the constant, the listed values are partial effects of setting the dummy variables to one. Left-out categories are: High incentive treatment, all payoffs non-negative, male, age 18-34, primary / lower secondary education, net annual household income below 22,000 Euros, total wealth below 10,000 Euros, not being the household's financial administrator, not being financially knowledgeable (self-rated).
† The low incentive treatment enters multiplicatively.

Table 4: Estimated Loss Aversion Parameters (λ)

	CentERpanel		Laboratory	
Covariate	**(4),(6)**	**(4),(6)**	**(4),(6)**	**(4),(6)**
Constant	2.38***	2.94***	1.67***	1.27
	(0.159)	(0.543)	(0.243)	(0.182)
Female		0.617*		0.766**
		(0.329)		(0.302)
Age 35-44		1.14**		
		(0.541)		
Age 45-54		-0.752*		
		(0.405)		
Age 55-64		-0.0463		
		(0.482)		
Age 65+		-0.673		
		(0.447)		
Hi Sec Educ / Int Voc Train		0.205		
		(0.391)		
Higher Voc Train		0.146		
		(0.435)		
University		0.663		
		(0.586)		
Income EUR 22k-40k		-0.378		
		(0.312)		
Income EUR 40k+		-1.15***		
		(0.388)		
Wealth EUR 10k-50k		0.193		
		(0.489)		
Wealth EUR 51k-200k		-0.0747		
		(0.367)		
Wealth EUR 201k+		-0.0697		
		(0.429)		
HH Financial Admin		-0.183		
		(0.299)		
Financially Knowledgeable		-0.418		
		(0.297)		
Hypothetical Treatment	1.31***	1.37***	0.0302	0.0418
	(0.317)	(0.441)	(0.382)	(0.243)
Low Incentive Treatment †	0.861***	0.839***	0.862**	0.934
	(0.0290)	(0.0433)	(0.0702)	(0.0529)

Note: Number of Observations is 1,422 (CentERpanel) and 178 (Laboratory), respectively. Estimation follows (11) based on the utility functions in the column headers. Regression coefficients are transformed back to the original scale by inverting the function on the left-hand side of (10). Except for the constant, the listed values are partial effects of setting the dummy variables to one. Left-out categories are: High incentive treatment, all payoffs non-negative, male, age 18-34, primary / lower secondary education, net annual household income below 22,000 Euros, total wealth below 10,000 Euros, not being the household's financial administrator, not being financially knowledgeable (self-rated).

† The low incentive treatment enters multiplicatively.

Table 5: Estimated Uncertainty Resolution Preferences (ρ)

	CentERpanel		**Laboratory**	
Covariate	**(4),(6)**	**(4),(6)**	**(4),(6)**	**(4),(6)**
Constant	1.01	1.04	0.931	0.891
	(0.0250)	(0.0849)	(0.0753)	(0.0812)
Female		0.0179		0.321***
		(0.0457)		(0.0620)
Age 35-44		-0.0054		
		(0.0686)		
Age 45-54		0.0069		
		(0.0694)		
Age 55-64		-0.134*		
		(0.0730)		
Age 65+		-0.207***		
		(0.0763)		
Hi Sec Educ / Int Voc Train		-0.0446		
		(0.0563)		
Higher Voc Train		-0.0903		
		(0.0620)		
University		-0.0100		
		(0.0811)		
Income EUR 22k-40k		0.0274		
		(0.0521)		
Income EUR 40k+		-0.0609		
		(0.0685)		
Wealth EUR 10k-50k		0.0895		
		(0.0744)		
Wealth EUR 51k-200k		0.0402		
		(0.0609)		
Wealth EUR 201k+		0.0444		
		(0.0742)		
HH Financial Admin		0.0588		
		(0.0517)		
Financially Knowledgeable		-0.0377		
		(0.0529)		
Hypothetical Treatment	-0.0964**	-0.109**	-0.177*	-0.189*
	(0.0405)	(0.0460)	(0.104)	(0.0988)
Low Incentive Treatment †	0.999	1.00	1.00	0.992
	(0.0015)	(0.0056)	(0.0167)	(0.0285)

Note: Number of Observations is 1,422 (CentERpanel) and 178 (Laboratory), respectively. Estimation follows (11) based on the utility functions in the column headers. Regression coefficients are transformed back to the original scale by inverting the function on the left-hand side of (10). Except for the constant, the listed values are partial effects of setting the dummy variables to one. Left-out categories are: High incentive treatment, all payoffs non-negative, male, age 18-34, primary / lower secondary education, net annual household income below 22,000 Euros, total wealth below 10,000 Euros, not being the household's financial administrator, not being financially knowledgeable (self-rated).

† The low incentive treatment enters multiplicatively.

Table 6: Estimated Random Choice Probabilities (ω)

	CentERpanel		Laboratory	
Covariate	**(4),(6)**	**(4),(6)**	**(4),(6)**	**(4),(6)**
Constant	0.0832***	0.102***	0.0049	0.0027
	(0.0083)	(0.0216)	(0.0053)	(0.0038)
Female		0.0197*		0.0034
		(0.0118)		(0.0045)
Age 35-44		0.0201		
		(0.0182)		
Age 45-54		0.0479**		
		(0.0203)		
Age 55-64		0.141***		
		(0.0326)		
Age 65+		0.293***		
		(0.0502)		
Hi Sec Educ / Int Voc Train		-0.0445***		
		(0.0133)		
Higher Voc Train		-0.0535***		
		(0.0140)		
University		-0.0622***		
		(0.0155)		
Income EUR 22k-40k		-0.0107		
		(0.0121)		
Income EUR 40k+		-0.0217		
		(0.0149)		
Wealth EUR 10k-50k		-0.0362**		
		(0.0144)		
Wealth EUR 51k-200k		-0.0249**		
		(0.0124)		
Wealth EUR 201k+		-0.0316**		
		(0.0134)		
HH Financial Admin		0.0005		
		(0.0116)		
Financially Knowledgeable		-0.0136		
		(0.0115)		
Short Duration		0.0822***		-0.0015
		(0.0207)		(0.0022)
Long Duration		-0.0485***		0.0032
		(0.0130)		(0.0064)
Hypothetical Treatment	0.0070	0.0049	0.0043	0.0017
	(0.0106)	(0.0127)	(0.0054)	(0.0036)
Low Incentive Treatment †	1.11***	1.06***	3.77***	2.74***
	(0.0078)	(0.0082)	(0.0150)	(0.0147)

Note: Number of Observations is 1,422 (CentERpanel) and 178 (Laboratory), respectively. Estimation follows (11) based on the utility functions in the column headers. Regression coefficients are transformed back to the original scale by inverting the function on the left-hand side of (10). Except for the constant, the listed values are partial effects of setting the dummy variables to one. Left-out categories are: High incentive treatment, all payoffs non-negative, male, age 18-34, primary / lower secondary education, net annual household income below 22,000 Euros, total wealth below 10,000 Euros, not being the household's financial administrator, not being financially knowledgeable (self-rated), completion time between 9 and 18 minutes.

† The low incentive treatment enters multiplicatively.

Table 7: Correlation Matrices, τ, and Log-Likelihoods

A: (4),(6)	**CentERp, few Covariates**			
	σ_γ	σ_λ	σ_ρ	σ_ω
σ_γ	0.037			
	(0.001)			
σ_λ		1.530		
		(0.042)		
σ_ρ			0.452	
			(0.024)	
σ_ω				1.957
				(0.090)
τ	4.072			
	(0.068)			
$\tau_{\text{Low Inc}}$ †	0.281			
	(0.008)			
Log-Likel	30234.4			

B: (4),(6)	**CentERp, all Covariates**			
	σ_γ	σ_λ	σ_ρ	σ_ω
σ_γ	0.037			
	(0.001)			
σ_λ		1.596		
		(0.049)		
σ_ρ			0.456	
			(0.025)	
σ_ω				1.811
				(0.084)
τ	4.011			
	(0.069)			
$\tau_{\text{Low Inc}}$ †	0.286			
	(0.008)			
Log-Likel	30079.6			

C: (4),(6)	**Lab, few Covariates**			
	σ_γ	σ_λ	σ_ρ	σ_ω
σ_γ	0.020			
	(0.001)			
σ_λ		0.959		
		(0.080)		
σ_ρ			0.415	
			(0.088)	
σ_ω				2.729
				(0.717)
τ	4.200			
	(0.131)			
$\tau_{\text{Low Inc}}$ †	0.255			
	(0.015)			
Log-Likel	3507.1			

D: (4),(6)	**Lab, all Covariates**			
	σ_γ	σ_λ	σ_ρ	σ_ω
σ_γ	0.020			
	(0.001)			
σ_λ		0.755		
		(0.081)		
σ_ρ			0.401	
			(0.075)	
σ_ω				2.858
				(0.864)
τ	4.278			
	(0.134)			
$\tau_{\text{Low Inc}}$ †	0.264			
	(0.015)			
Log-Likel	3496.9			

Note: Number of Observations is 1,422 (CentERpanel) and 178 (Laboratory), respectively. Estimation follows (11) and the utility function specified in the top corners of the individual panels. The correlation matrix contains standard deviations of the untransformed normal distribution on the diagonal.
† The low incentive treatment enters multiplicatively.

D Figures

Figure 1: Screenshot of Sheet 5, First Screen

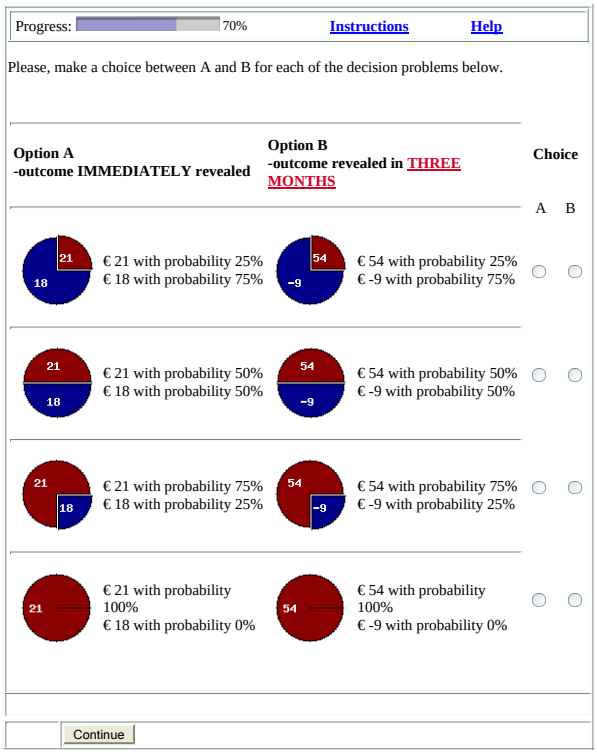


Figure 2: Choices Made by Selected Individuals

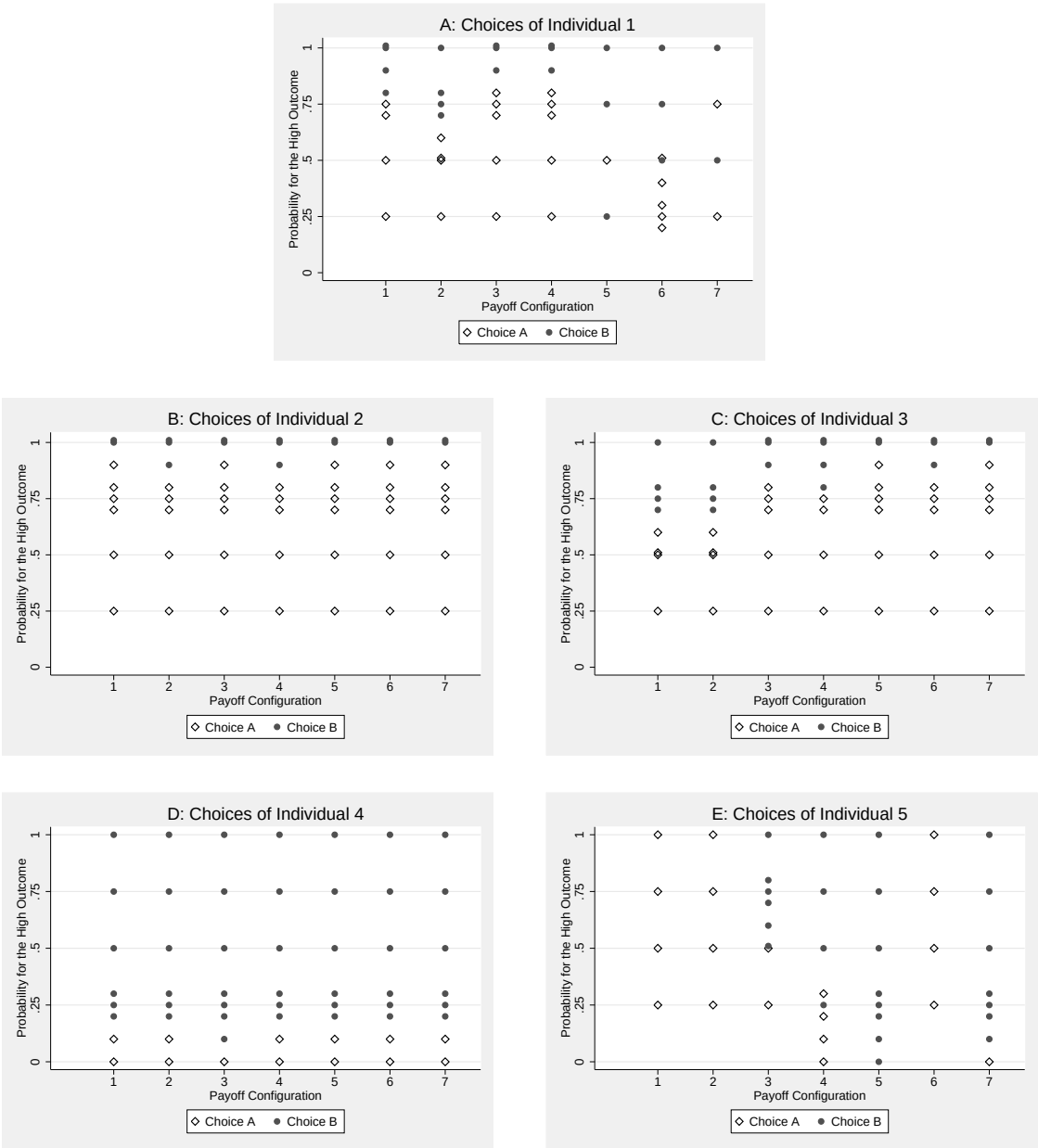
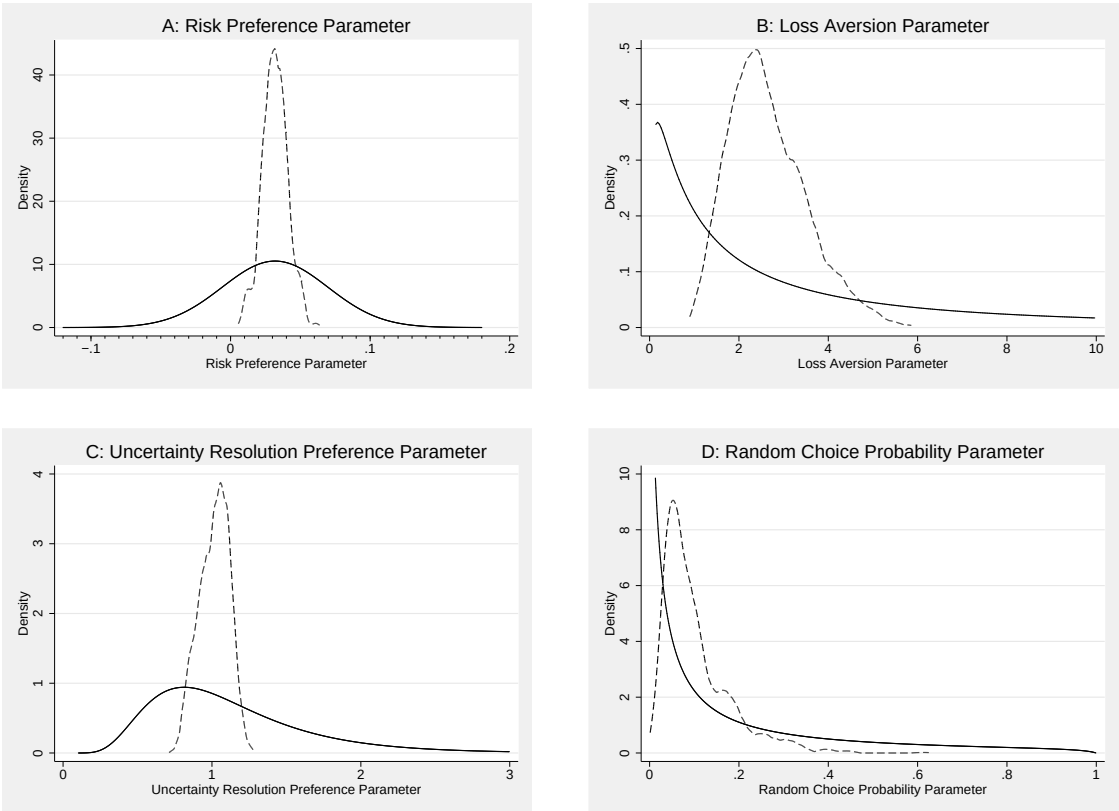
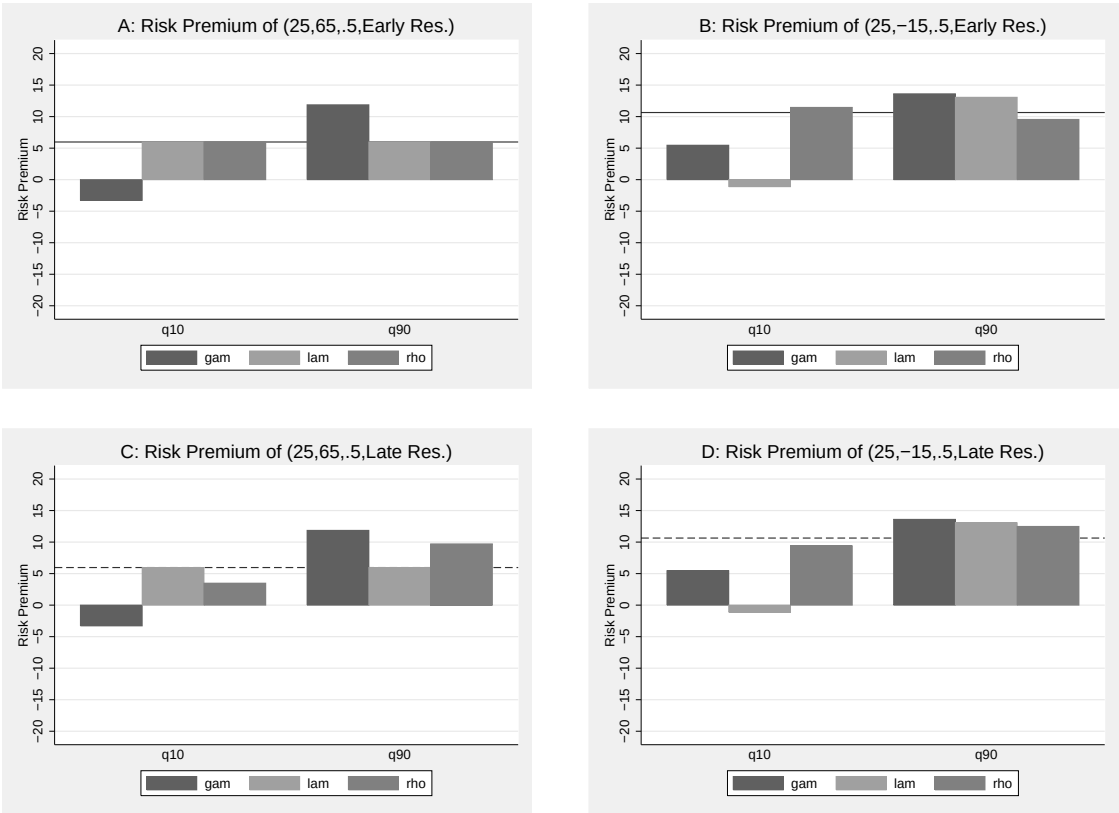


Figure 3: The Distributions of Preference and Error Parameters in the Population: Comparing Total Heterogeneity and Observed Heterogeneity



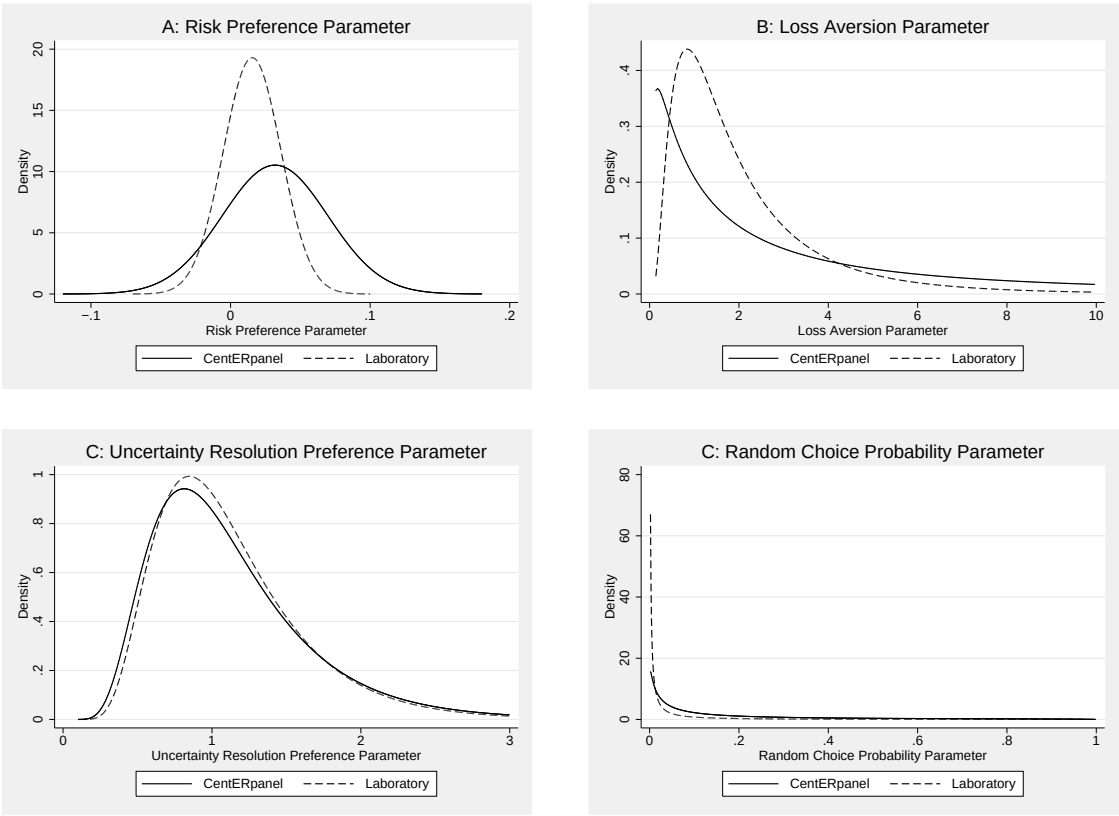
Note: Solid lines are estimated parameter distributions taking observed and unobserved heterogeneity into account. Dashed lines neglect the unobserved part, they are kernel density estimates over the socio-demographic group means. Both are based on the model accounting for all covariates (second columns of Tables 3–7). Treatment effects are netted out.

Figure 4: Risk Premia by Preference Parameter Quantiles



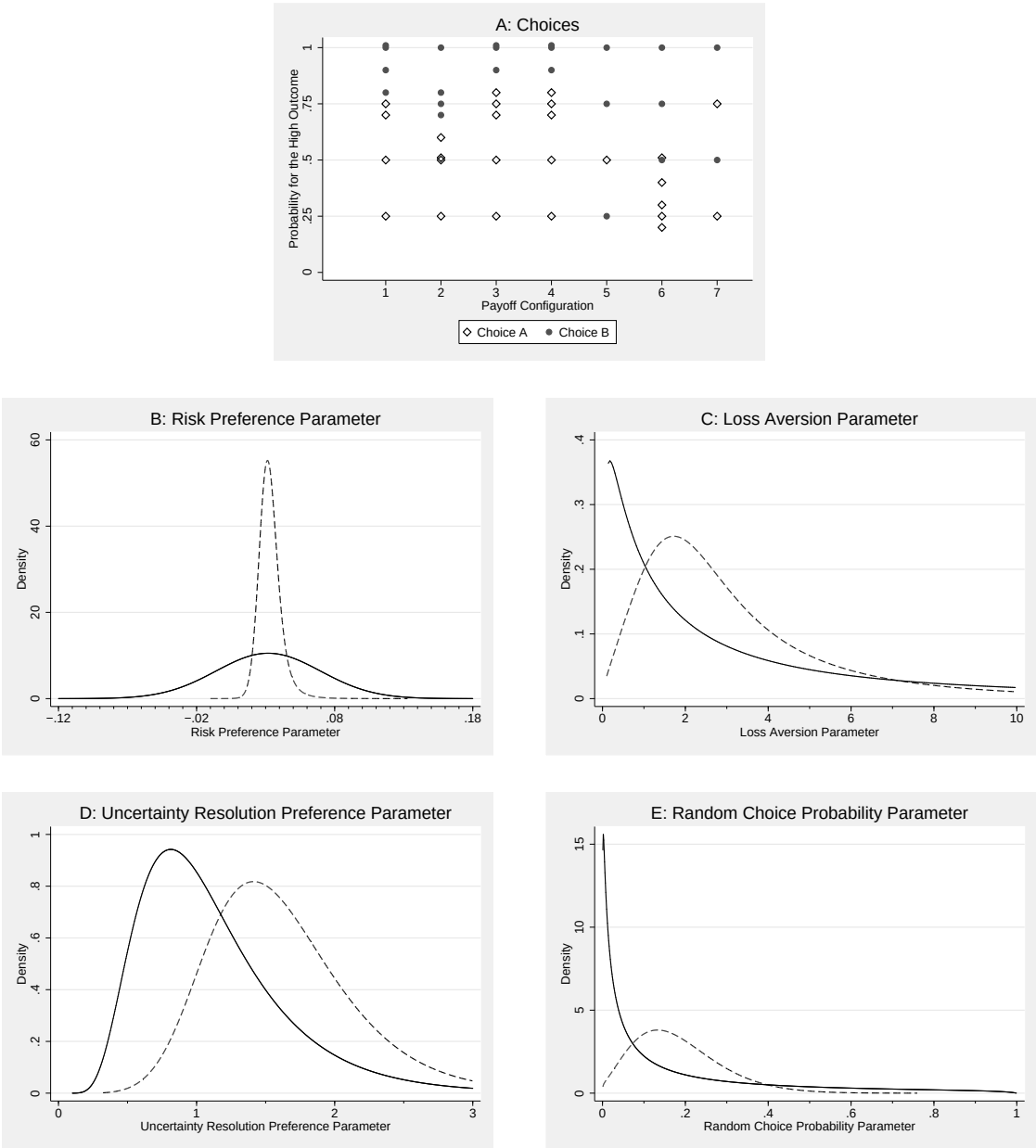
Note: Solid (dashed) lines depict the risk premia for the early (late) resolving lottery, evaluated at the median parameter estimates. The bars depict the risk premia when setting the parameters one at a time to their 10% and 90% quantiles.

Figure 5: Comparing the Distributions of the Preference and Error Parameters from the CentERpanel and Laboratory Experiments



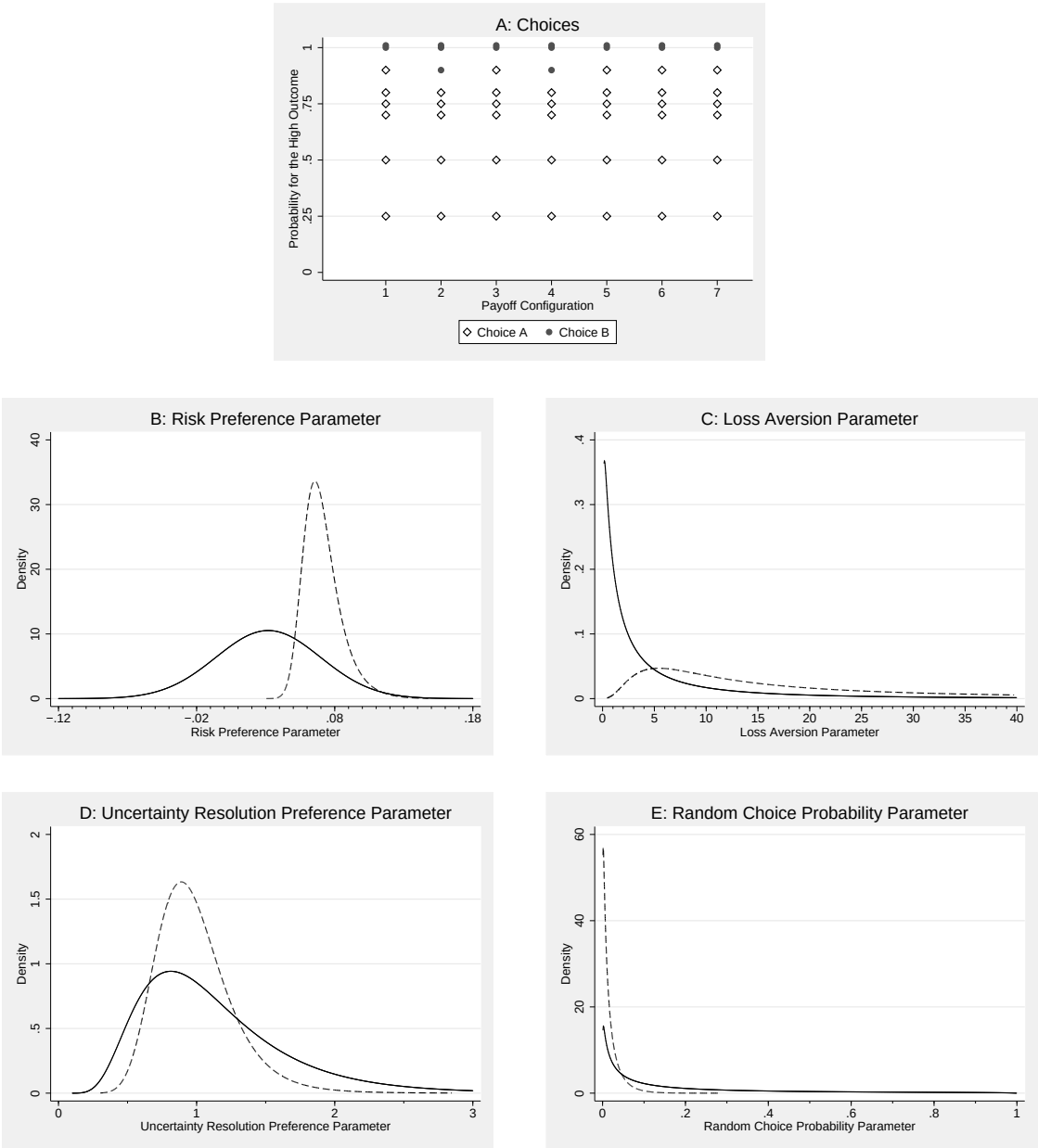
Note: Both lines are estimated parameter distributions taking observed and unobserved heterogeneity into account. They are based on the models accounting for all available covariates (second and fourth columns of Tables 3–7). Treatment effects are netted out.

Figure 6: Choices and Preference Parameter Distributions of Individual 1



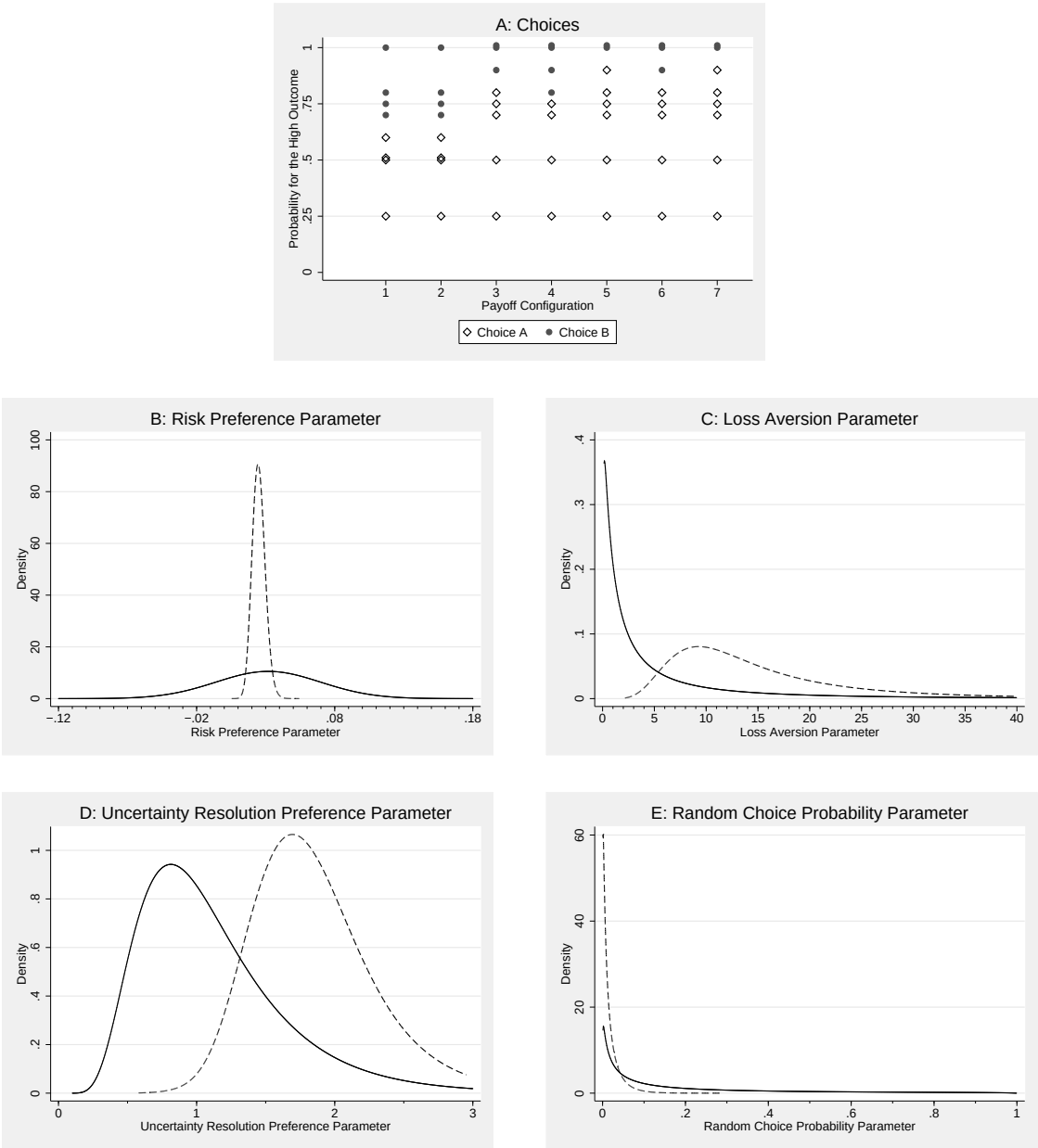
Note: Solid lines are the estimated parameter distributions for individual 1's socio-demographic group (female, age 45-54, higher secondary education or intermediate vocational training, household income between 22,000 Euros and 40,000 Euros, wealth above than 200,000 Euros, not financially knowledgeable, financial administrator, medium duration). Dashed lines are the marginal distributions of parameters conditional on the choices shown in the first panel. Graphs are based on estimates in Tables 3–7, second columns / panel.

Figure 7: Choices and Preference Parameter Distributions of Individual 2



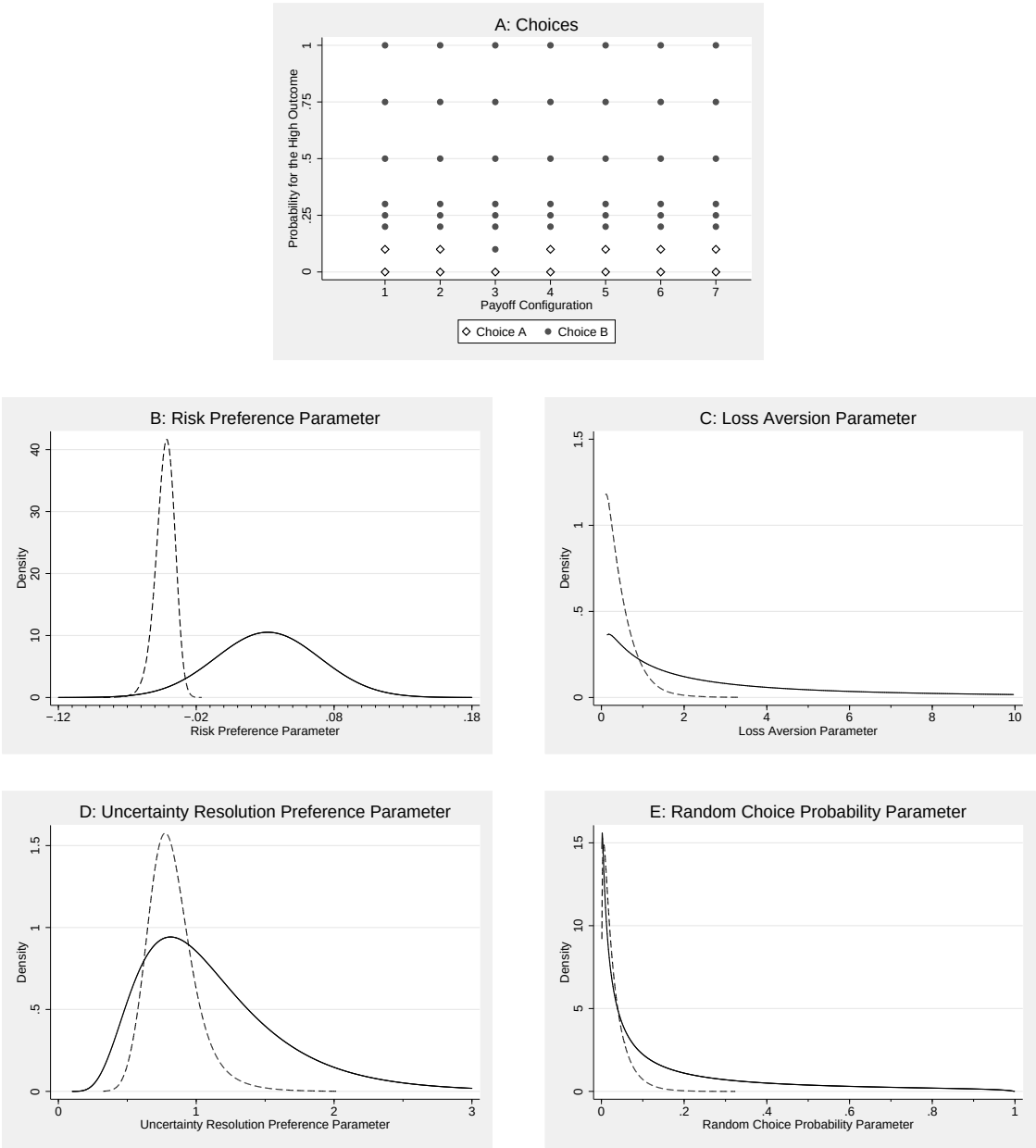
Note: Solid lines are the estimated parameter distributions for individual 2's socio-demographic group (female, age 35-44, higher vocational training, household income between 22,000 Euros and 40,000 Euros, wealth above 200,000 Euros, not financially knowledgeable, not the financial administrator, medium duration). Dashed lines are the marginal distributions of parameters conditional on the choices shown in the first panel. Graphs are based on estimates in Tables 3–7, second columns / panel.

Figure 8: Choices and Preference Parameter Distributions of Individual 3



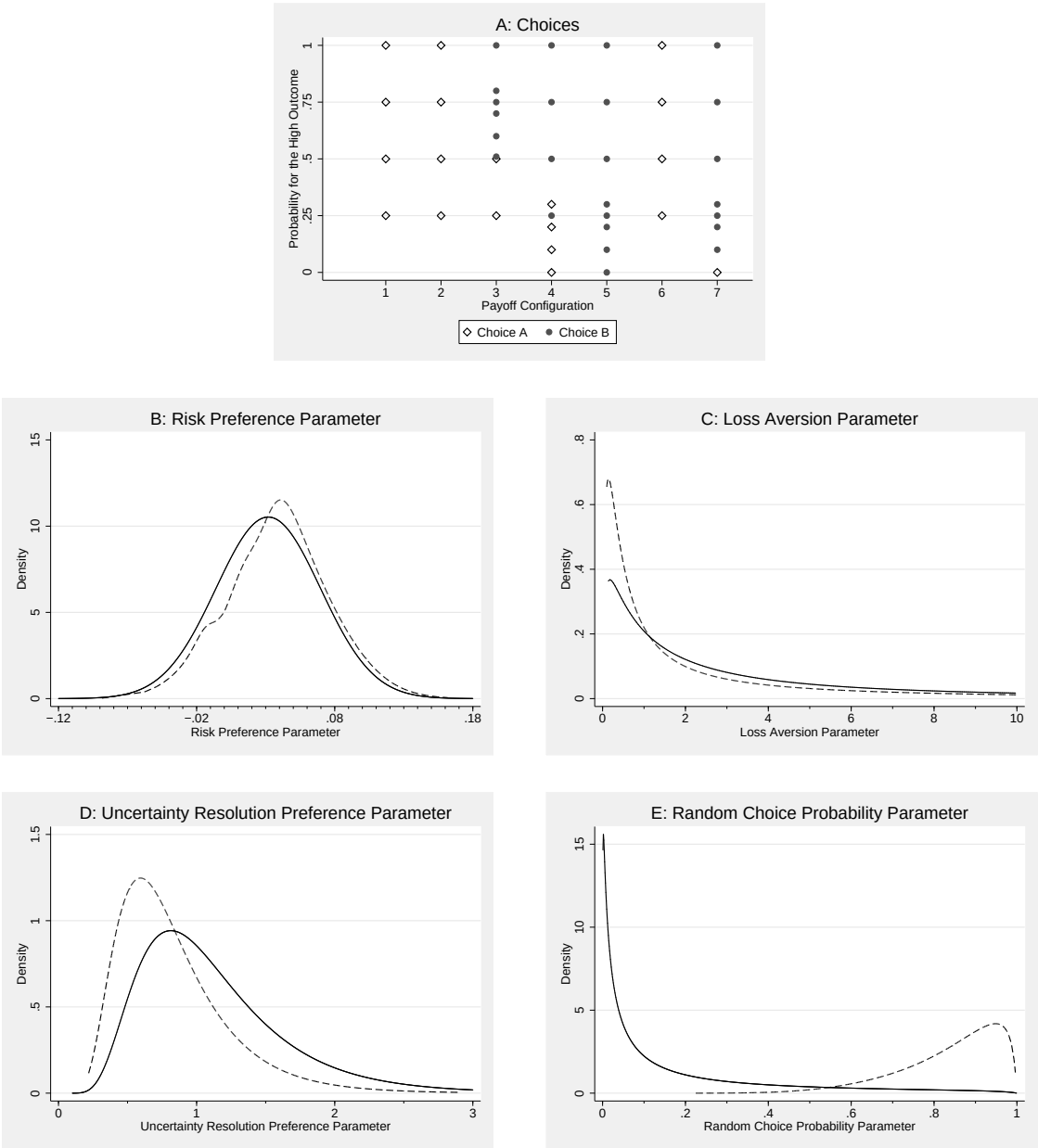
Note: Solid lines are the estimated parameter distributions for individual 3's socio-demographic group (female, age 18-34, higher secondary education or intermediate vocational training, household income below 22,000 Euros, wealth between 10,000 Euros and 50,000 Euros, financially knowledgeable, financial administrator, medium duration). Dashed lines are the marginal distributions of parameters conditional on the choices shown in the first panel. Graphs are based on estimates in Tables 3–7, second columns / panel.

Figure 9: Choices and Preference Parameter Distributions of Individual 4



Note: Solid lines are the estimated parameter distributions for individual 4's socio-demographic group (male, age 45-54, primary or lower secondary education, household income above 40,000 Euros, wealth between 51,000 Euros and 200,000 Euros, not financially knowledgeable, financial administrator, short duration). Dashed lines are the marginal distributions of parameters conditional on the choices shown in the first panel. Graphs are based on estimates in Tables 3–7, second columns / panel.

Figure 10: Choices and Preference Parameter Distributions of Individual 5



Note: Solid lines are the estimated parameter distributions for individual 5's socio-demographic group (female, aged at least 65, primary or lower secondary education, household income below 22,000 Euros, wealth above 200,000 Euros, not financially knowledgeable, financial administrator, medium duration). Dashed lines are the marginal distributions of parameters conditional on the choices shown in the first panel. Graphs are based on estimates in Tables 3–7, second columns / panel.