Beliefs in Reciprocity, Confidence, and Trust*

Mohammed Abdellaoui[1], Levent Gumus[2], Yassine Kaouane[3], Emmanuel Kemel[1], and Ferdinand M. Vieider[2,3]

[1]*CNRS and HEC Paris, France*
[2]*RISL$\alpha\beta$, Department of Economics, Ghent University, Belgium*
[3]*RISL$\alpha\beta$ Africa, AIRESS, University Mohammed VI Polytechnic, Morocco*

March 8, 2024

Abstract

Probabilistic beliefs are key drivers of economic behaviour, yet incentive-compatible and easy-to-understand measures of beliefs have proved elusive. We develop a novel method that allows us to econometrically recover probabilistic beliefs from simple binary choices between bets on different events. The method is incentive-compatible under all major models of decision-making, and easy to understand for subjects. We test the method by eliciting beliefs about the reciprocation of trust. The mean belief we elicit predicts trust, and performs slightly better than commonly used qualitative and quantitative survey questions. Further adding confidence in beliefs describing the belief *distribution* improves the performance of our measure dramatically. This shows the promise of the method for applied work.

1 Introduction

Probabilistic beliefs about outcome-generating events play a key role for most any economic decision, from investment choices to strategic interactions. Such beliefs have, however, proved challenging to quantify by means of simple, incentive-compatible tasks. While a variety of belief measures have been used especially in the literature on interactive games, such measures are often introspective (Manski, 2004; Bellemare, Kröger and Van Soest, 2008). Revealed-preference measures of beliefs such as scoring rules have been shown to suffer from failures of incentive-compatibility when subjects have preferences deviating from expected value maximization. Generalizations of these measures devised to make them incentive-compatible, such as the one proposed by Hossain and Okui (2013), are difficult to understand for experimental subjects, thus casting doubt on the behavioural reactions they may trigger (Danz, Vesterlund and Wilson,

*The present work has received IRB approval from the ethics committee at the Faculty of Business and Economics at Ghent University under number FEB2020G, as well as from University Mohammed VI Polytechnic. Ferdinand Vieider and Levent Gumus gratefully acknowledge financial support by the Research Foundation Flanders under Grant Number G008421N.

2022; see also Schotter and Trevino, 2014, for a detailed discussion of the issues). Heinemann, Nagel and Ockenfels (2009) and Offerman, Sonnemans, van de Kuilen and Wakker (2009) have proposed methods to correct scoring rules ex post with additional preference measurements, but such approaches may be sensitive to measurement noise. Other revealed-preference methods eliciting matching probabilities or cash equivalents need strong assumptions about preferences (Baillon, Bleichrodt, Keskin, L'Haridon and Li, 2018a), and may be subject to possible error propagation (Abdellaoui, Bleichrodt and Gutierrez, 2024).

Here we develop and test an incentive-compatible method to measure subjective belief distributions. Probabilistic beliefs are inferred from a series of binary choices between exogenously determined events to win a fixed prize. Specifically, we use these binary choices to econometrically estimate parametric subjective probability distributions. The method is robust to both probability distortions, one of the main sources of descriptive discrepancies from the standard model of rational choice under uncertainty, and utility curvature. The method is much simpler to administer and understand than methods based on choice lists systematically changing cutoffs between events (Abdellaoui, Bleichrodt, Kemel and L'Haridon, 2021), and is likely to reduce noise relative to the latter (see Bouchouicha, Wu and Vieider, 2023, for evidence of noise arising from choice list designs under risk). The main cost of these advantages is that the recovery of beliefs requires the use of a parametric belief distribution. The latter can however easily be found for most applications, so that we see this as a relatively minor limitation.

We test our method by measuring beliefs about the probability of reciprocation in a trust game. Beliefs about trustworthiness play a central role in trusting decisions (Fehr, 2009). We thus let subjects play a variation on the trust game before measuring their probabilistic beliefs about the likelihood of reciprocity.[1] The latter are elicited by letting subjects bet on the proportion of reciprocating individuals in a large group of experimental subjects. This setting has several advantages. Having individual measures of trust allows us to assess the predictive power of our belief measures for trusting behaviour, which constitutes a test for our belief-measurement method. The explicitly frequentist framing makes it easy for subjects to understand how uncertainty will be resolved. A Beta distribution seems a natural choice for the bets on the proportion of reciprocators in a population we consider, and it is advantageous because of its great flexibility.

[1] We developed the variation of the trust game specifically to avoid conditionality of reciprocating decisions on initial offers. This allows us to obtain a continuous measure of trust, which gives us sufficient power in regression analysis, while keeping the offer to the second-mover constant. The reciprocating decision is thus decoupled from the level of trust displayed by the first mover, which in turn allows for a straightforward measurement of beliefs about reciprocation. See the description of the experiment for details.

The mean of the Beta reflects a subject's expectation about the percentage of reciprocators. The concentration of the Beta reveals the amount of uncertainty the subject perceives as surrounding the expected percentage of reciprocators. It consequently measures the confidence the subject has in its assessment. The use of the Beta concentration echoes the idea of a second-order probability distribution over the set of possible percentages of reciprocators, i.e., the set of priors (Klibanoff, Marinacci and Mukerji, 2005).

We use a fairly minimal set of binary choices to obtain our belief measurements, consisting of 39 choices plus 5 repetitions. Belief measurements take some 6-10 minutes including instructions, showing the viability of the method in cases where beliefs are to be used as an explanatory variable. The flip side is that identification of individual-level parameters relies on relatively few data points. We take the uncertainty arising from this seriously by using a Bayesian hierarchical approach to analyze the data. This approach yields an optimal compromise between aggregate and individual-level data, weighing individual-level estimates by their relative reliability. Estimated belief parameters furthermore have a probabilistic interpretation, given that parameters are uncertain quantities in Bayesian statistics, while the data are considered given (in the true sense of the Latin *data*, for *given*). We regress our measure of trust on the parameters capturing the mean belief and confidence in the mean belief in the same model in which we estimate these parameters, thus taking the full posterior uncertainty about these parameters into account. We describe the method in some detail and make the code publicly available, to provide a toolbox for applied researchers interested in belief measurement.

We find considerable heterogeneity in mean beliefs, as well as in confidence about beliefs as measured by the concentration parameter of the Beta distribution.[2] Our measures of beliefs about reciprocity have significant explanatory power for trusting behaviour. The mean belief explains about 7% of variation in trust. We compare this to several commonly used survey measures. At between 4% and 5% of variance explained, a more fine-grained variation of the World Value Survey measure and an introspective quantitative measure simply asking for the proportion of reciprocators perform best, although they still fall short of our incentivized measure. We next examine whether our measure of confidence in beliefs has any explanatory power above the mean belief. We find that confidence in beliefs has a strongly positive effect on trust. Conditional on the inclusion of mean beliefs, confidence in beliefs is indeed the single most important explanatory variable, bringing the explained variance to over 20%. Further adding a survey question on the

[2]We mainly use the concentration because of its intuitive interpretation and because it is orthogonal to our measure of the average belief, but all the results are stable to using alternative measures of dispersion such as the variance instead.

'general willingness to take risk' and controls for age and gender brings this figure to 24%—a remarkable figure compared to typical figures in the decision-making literature (von Gaudecker, van Soest and Wengström, 2011; L'Haridon and Vieider, 2019; Di Falco and Vieider, 2022).[3]

Our findings also add to an extensive literature discussing the importance of beliefs about reciprocity on trust, as well as adding to the literature on the validity of survey measures of beliefs. Such survey measures have been used extensively in applied work (Knack and Keefer, 1997; Tabellini, 2010; Rohner, Thoenig and Zilibotti, 2013), but their validity has at times been disputed (Glaeser, Laibson, Scheinkman and Soutter, 2000). While we also find a correlation of trust with measures of risk taking—which is clearly relevant theoretically (Karlan, 2005; Fehr, 2009), though empirically its influence is disputed (Bohnet and Zeckhauser, 2004; Houser, Schunk and Winter, 2010; Li, Turmunkh and Wakker, 2020)—our measures of preferences are rather limited. Li et al. (2019) explore the impact of ambiguity attitudes on trusting behaviour, using a nonparameteric measurement method proposed by Baillon, Huang, Selim and Wakker (2018b) to obtain attitudes towards natural sources of uncertainty without the need to first measure beliefs. While their preference measures are thus much richer than ours, they only show the effect of the difference in point beliefs about the best and worst of three events. Engelmann, Friedrichsen, van Veldhuizen, Vorjohann and Winter (2023) propose a decomposition of the trust game, and measure beliefs by an introspective question incentivized with a linear scoring rule. They also use a matching probability jointly capturing beliefs and attitudes towards trust. Both measures only provide point estimates. Our approach is thus highly complementary to these approaches by specifically focusing on the impact of more sophisticated measures of probabilistic beliefs.

2 Study Design

2.1 General setup and implementation

We ran the experiment at the laboratory of University Mohammed VI Polytechnic in Rabat, Morocco. 127 subjects (75% female) took part in the experiment between October and December 2022. All the participants were students at the Faculty of Governance, Economics and Social Sciences. They were recruited through posters and an email sent to all students. The computer-

[3] A comparison to the explanatory power of other measures of beliefs is difficult, given that the power of a belief measure in explaining behaviour will depend crucially on the type of task used. In particular, the trust game we study here could potentially be affected by risk attitudes (Karlan, 2005; Schechter, 2007), ambiguity attitudes (Li, Turmunkh and Wakker, 2019), or betrayal aversion (Bohnet and Zeckhauser, 2004), amongst others. We thus mostly benchmark the performance internally against commonly used qualitative and quantitative introspective measures, which have been used widely in the literature.

based experiment was conducted through individual interviews of three subjects at a time, using software specifically developed for the experiment. Each participant was seated in front of a screen in the presence of the experimenter. We used small sessions to improve the quality of the data. No communication between participants was allowed during the experiment. At their arrival, subjects were presented with video instructions, after which they were asked to sign the consent form.[4] This was followed by a series of comprehension questions and a set of practice questions so that the participants could get familiar with the software. No subject refused to participate or was excluded. The experiment took about 30 minutes in total.

The experiment was run using the strategy method, i.e. each subject played the role of first mover or trustor, and of second mover, or trustee. The payoff-relevant role was randomly determined at the end. A pilot experiment showed that it made no difference whether the role of trustor and trustee were assigned between subjects or within subjects (results available upon request). Having subjects play both roles has the further advantage of allowing us to examine how beliefs relate to a subject's own reciprocating decision. All subjects sequentially participated in three tasks: 1) a task to elicit 'trust equivalents'; 2) the trustee's decision on whether to split the pie equally or unequally; and 3) a series of binary choices to elicit revealed beliefs about reciprocation. Below, we describe each part in turn.

2.2 Trust equivalents

We use trust equivalents to measure trust—a version of the trust game fashioned after the binarized version of Bohnet and Zeckhauser (2004) and Trautmann, van de Kuilen and Zeckhauser (2013). The game is shown in figure 1. Binarizing the response has the advantage that reciprocators only need to take one simple decision, and that the monetary consequence of that decision are clear and transparent. Equalizing the outcomes of the trustor and trustees eliminates concerns for others' outcomes and reduces efficiency concerns. Other than in the binary choice version of the game, however, we determine the sure amount paid to both players that makes the decision-maker indifferent between delegating the decision of splitting the pie to the trustee, and accepting a sure amount for both players. Determining a *trust equivalent* (*TE*) in this way gives us a rich measure of trust, while eliminating potential endowment effects that could otherwise introduce additional heterogeneity in behaviour.

The trust equivalent was always elicited in the first part of the experiment. This allowed us to

[4] The video instructions are available online. First video for the trust game: https://youtu.be/56aBLsM0Cpghttps. Second video for the beliefs elicitation: https://youtu.be/M7ys63m8TfM

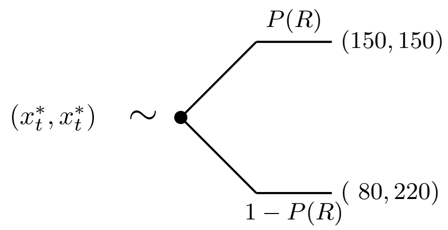

Figure 1: Trust equivalents

fully describe the trust game, and was meant to reduce contamination effects from reciprocating decisions. The outcomes, denominated in Moroccan Dirhams (*Dhs*), constitute significant sums of money for our participants (at the time of the experiment, Dhs 10 corresponded approximately to €1). The sure outcome x changed between Dhs 80 and Dhs 150 in steps of Dhs 5. To help subjects converge to the point of indifference, a bisection procedure was implemented starting from a randomly selected point. Subjects could thus focus on a simple binary choice at a time. After 5 choices, a point of indifference was reached.

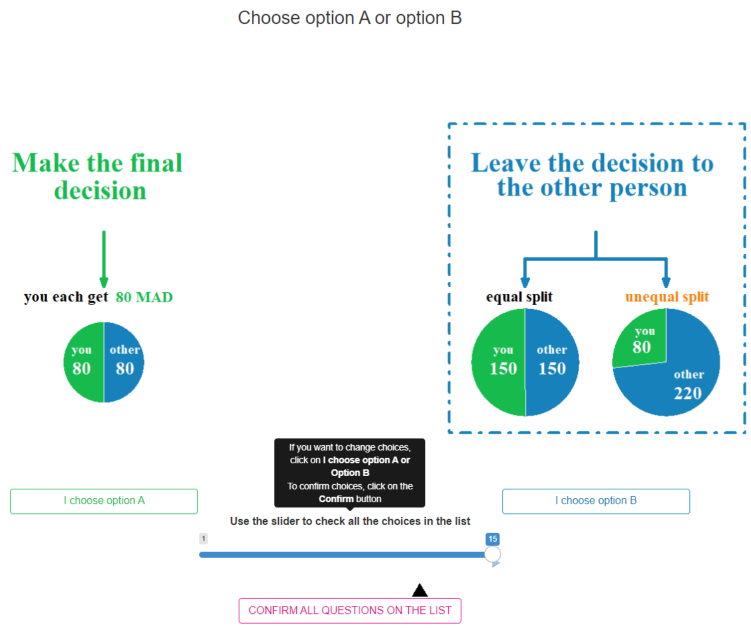

Figure 2: Slider to adjust choices after bisection procedure

To avoid issues with incentive-compatibility arising from the use of bisection methods, subjects were told that this procedure served purely the purpose of helping them to fill in an underlying list, based on which they would be paid. Once the list had been filled in, they were able to revisit their choices, and to adjust them in case they did not feel comfortable with them. A screenshot of the procedure is shown in figure 2. Once they had reviewed their choices and potentially adjusted their preference, they could confirm their choice and move on to the following task.

2.3 The trustee's decision

After responding to the trust equivalent question, each participant took on the role of trustee and chose between two possibilities on how to split the pie. Subjects were informed that the role counting for the actual payment would be randomly assigned in the end, and that their decision could thus determine their own payoff, as well as the payoff of another person in the role of the trustor (neutrally called 'player 1' in the instructions), with whom they would be randomly paired. Given that the game had already been explained, the explanations for this part were fairly short due to the binary nature of the game. Subjects learned that they had been entrusted with Dhs 300, and that they got to make the final decision on how to split that money. The two options were an 'equal split', with Dhs 150 being paid both to themselves and to the other person with whom they would be paired; and an 'unequal split', whereby they could split the amount in such a way as to give Dhs 80 to the person they were paired with, while attributing Dhs 220 to themselves.

2.4 Eliciting beliefs about the proportion of reciprocators

Our main interest is the subjective probability distribution subjects hold over the proportion of reciprocators in the population. To make the problem as intuitive as possible, we formulated all choices based on the number of people out of 100 choosing an equal split of the pie in step 2 above. We developed a binary choice procedure that allows us to collect information on which of two events is considered more likely. This emulates methods based on the idea of exchangeable events used in Baillon (2008) and Abdellaoui, Baillon, Placido and Wakker (2011). Being based entirely on independent binary choices, however, the method circumvents the issues of incentive compatibility and error propagation incurred by the latter.

All bets are placed to obtain a fixed prize of Dhs 100 or else 0. A bet on an uncertain event E consequently corresponds to the gamble that pays 100 if the event occurs and nothing otherwise, which we denote 100_E0. The use of such bets serves to keep the task simple for subjects, who could thus focus on the cutoff points of the two events since everything else was constant. In our experiment, events are generated by the same source of uncertainty: the number of subjects out of 100 splitting the pie equally. We assume that the value of the gamble 100_E0 is given by

$$\pi_E u(100) + (1 - \pi_E)u(0),$$

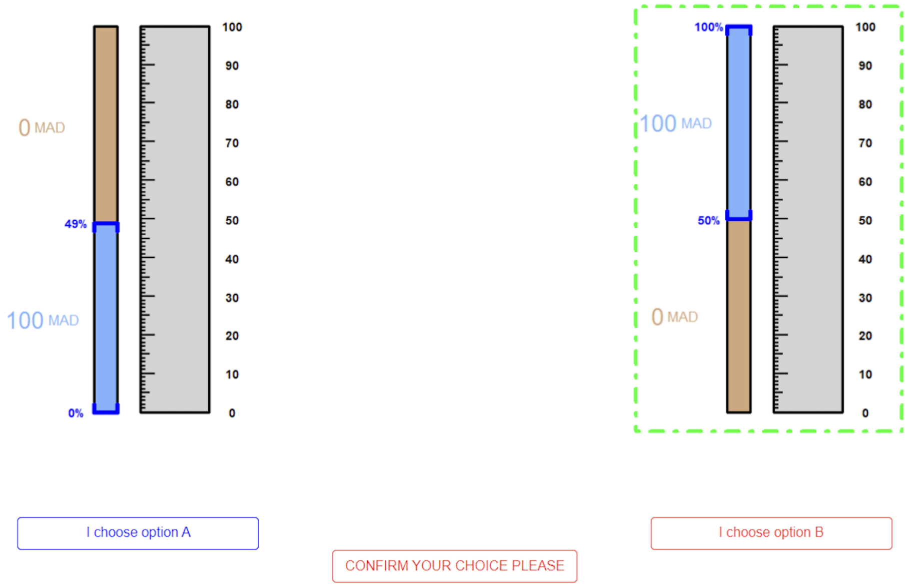

Figure 3: Screenshot of belief elicitation task

where π_E stands for the decision maker's willingness to bet on event E, and u stands for the utility function. This general model intersects a series of non-expected utility models that do not postulate that willingness to bet π_E coincides with the subjective probability of event E (multiple prior models as in Gilboa and Schmeidler, 1989, Ghiradato et al., 2003; rank-dependent utility as in Wakker 2010, and models based on probability intervals as in Manski, 2004).

We assume that there exists a subjective probability measure $P(.)$ on the state space, (the proportion of subjects splitting the pie equally) that holds without committing to expected utility maximization (Machina and Schmeidler, 1992; Trautmann and van de Kuilen, 2015). Combining these assumptions means that there exists a distortion function f such that, for any event E, $\pi_E = f(P(E))$. As a consequence, the gamble 100_A0 is preferred to 100_B0 implies that the subjective probability of A is at least as large as that of B, i.e., $P(A) \geq P(B)$.[5]

In other words, we use outright choice between bets generated by the same source of uncertainty to elicit subjective probabilities. This reduces to circumventing the effects of the curvature of f and u. Our method is cognitively simpler than methods based on exchangeability (Baillon, 2008; Abdellaoui et al., 2011; 2024), and it is further not affected by error propagation. Unlike previous methods based on exchangeability, our outright choices involve events that are not necessarily adjacent, which allows to compare "tail events" laying at the two extreme sides of the

[5]Specifically, preferring 100_A0 to 100_B0 means that $\pi_A \geq \pi_B$. As f is strictly increasing from $[0, 1]$ into $[0, 1]$, $P(A)(B)$ results.

domain. Eventually, our method circumvents the complexity of using risk to calibrate subjective probabilities as in the matching probability method (e.g., Abdellaoui et al., 2024).
[6]

We selected the choice tasks based on simulations such as to allow us to recover a wide range of mean and concentration parameters that might characterize subjects' beliefs. The tasks comprised the comparison of bets on events partitioning the whole state space (e.g., betting on less than 50 reciprocators versus 50 or more out of 100), the comparison of bets on "tail" events (e.g., betting on between 0 and 20 reciprocators, versus between 80 and 100), as well as betting on "internal" events versus their complement (e.g., a bet on there being between 20 and 50 reciprocators inclusive, versus the complement, made up of less than 20 or more than 50). We used a total of 39 unique binary choice tasks—Table 3 in Appendix B.1 shows the full list of choices. 5 randomly selected tasks were furthermore repeated, to allow for the better quantification of errors.

2.5 Econometric Recovery of Beliefs

We assume that beliefs are captured by a Beta distribution which appears a natural choice for the proportions we investigate, and which provides much flexibility in capturing different belief patterns. Let R be the proportion of reciprocators in the group. We can then write the probabilistic belief of a subject i about R as follows:

$$P_i(R \mid \alpha_i, \beta_i) = \mathcal{B}(\alpha_i, \beta_i). \tag{1}$$

For computational convenience, we reparametrize the Beta using its mean, $\mu_i = \frac{\alpha_i}{\alpha_i+\beta_i}$, and its concentration $\kappa_i = \alpha_i + \beta_i$. This is usually helpful in estimations, as μ_i and κ_i capture orthogonal dimensions. Consistent with the interpretation of κ_i in terms of confidence, this parameter can also be considered as revealing the cumulative experience a decision-maker has in similar situations (Aydogan, Baillon, Kemel and Li (2022)), with confidence thus increasing with past experience in similar situations. The shape parameters then obtain as $\alpha_i = \mu_i \, \kappa_i$ and $\beta_i = (1 - \mu_i) \, \kappa_i$.[7]

[6] For instance, Trautmann and van de Kuilen (2015) use a version of the matching probability method that assumes that the above model holds for both risk and uncertainty to infer the subjective probability of E from a matching probability, i.e., the objective probability p that makes the agent indifferent between 100_E0 and 100_p0.

[7] An alternative measure of dispersion one could use is the variance of the Beta distribution, $\nu^2 = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$. Such a measure has an interesting interpretation especially when applied to the performance of financial assets. It does, however, have the drawback the variance will be correlated with the mean for mechanical reasons, thus being less convenient in estimations. That being said, all the effect we present in this paper remain qualitatively

The Beta describing the subjective probability distribution over the proportion of reciprocators can now be used to predicts how a subject will choose between any given pair of events. In practice, we will invert this process by using the choices between events to estimate the unknown distribution governing a subject's choices. Let $P_i(E_j^A)$ and $P_i(E_j^B)$ denote the subjective probabilities of the events entailed by *"Option A"* and *"Option B"* in the j^{th} task. We can define these probabilities as follows

$$P_i(E_j^A) = \mathcal{F}(u_A \,|\, \mu_i\kappa_i, (1-\mu_i)\kappa_i) - \mathcal{F}(\ell_A \,|\, \mu_i\kappa_i, (1-\mu_i)\kappa_i)$$
$$P_i(E_j^B) = \mathcal{F}(u_B \,|\, \mu_i\kappa_i, (1-\mu_i)\kappa_i) - \mathcal{F}(\ell_B \,|\, \mu_i\kappa_i, (1-\mu_i)\kappa_i),$$

where $\mathcal{F}$ designates the cumulative distribution function of the Beta distribution in (1), and u and ℓ stand for the upper and lower bounds of the events in option A and option B, as indicated by the respective subscripts. This shows how the subjective probabilities capture the probability mass attributed to each given event by the underlying Beta.[8]

Given this setup, subject i will choose Option A over Option B whenever $P_i(E_j^A) \geq (P_i(E_j^B)$, or else will display the opposite choice pattern. We thus estimate the probability of subject i choosing Option A over Option B in choice j using the following logistic regression:

$$Pr_i^j[A \succ B] = \text{Bern}\left[\text{logit}^{-1}\left(\frac{P_i(E_{A_j}) - P_i(E_{B_j})}{\sqrt{2}\sigma}\right)\right], \tag{2}$$

where $Bern$ indicates the Bernoulli distribution, and $logit^{-1}$ the inverse-logit link function. The scaling by $\sqrt{2}$ serves merely to standardize the variance (Train, 2009).

We use a fairly minimal set of binary choices to elicit beliefs to showcase the viability of our method for applied work. To deal with the resulting uncertainty surrounding estimated parameter values, we use a random-parameter or *hierarchical* setup in which individual-level parameters are modelled as being drawn from an overarching distribution describing the entire experiment. Let $\boldsymbol{\theta}_i = \{logit^{-1}(\mu_i), log(\kappa_i)\}$ be a vector of individual-level parameters, where the inverse-logit and log transformations used serve to enforce the constraints arising from the Beta distribution. We can model the individual-level parameters as being drawn from a distribution

stable if we use the variance instead of the concentration of the Beta distribution.

[8] In some tasks, subjects could earn a prize based on a the union of two disjoined events, e.g. based on their being either less than 20 or more than 60 reciprocators. In those cases, we formally estimate the probability over the single, joint event first (e.g., there being between 20 and 60 reciprocators inclusive in the example above), and then define the winning probability as its complementary probability. See online appendix A for details and the annotated Stan code used in our estimations.

describing the subjects in the experiment as follows:

$$\boldsymbol{\theta}_i \sim m\mathcal{N}(\widehat{\boldsymbol{\theta}}, \Sigma), \tag{3}$$

where $m\mathcal{N}$ designates the multi-normal distribution, $\widehat{\boldsymbol{\theta}} = \{\widehat{\mu}, \widehat{\kappa}\}$ is a vector of aggregate means of the parameters, and Σ is a covariance matrix of the parameters. The aggregate parameters will thus act as endogenously-estimated priors for the individual-level parameters. While we need to specify hyperpriors for the aggregate parameters, which choose these to be one order of magnitude wider than the effects we would expect based on the scale of the data, thus assuring that they will not sway our results (a robustness analysis confirmed that this is indeed the case).

The hierarchical setup we use has the advantage that individual-level estimates will be weighed by their endogenously estimated reliability. That can be done because model parameters are themselves uncertain quantities in the Bayesian setup we use. Take e.g. μ_i^{ML}, the estimated mean belief for subject i in a maximum likelihood estimation. The estimate for μ_i^{ML} will typically follow an approximately normal distribution, so that we can model:

$$\begin{aligned} \mu_i^{ML} &\sim \mathcal{N}(\mu_i, \tau_i^2), \\ \mu_i &\sim \mathcal{N}(\widehat{\mu}, \Sigma_{11}), \end{aligned}$$

where τ_i captures the uncertainty surrounding to the maximum likelihood estimate of the mean belief (i.e. the standard error of the maximum likelihood estimate), μ_i is as usual the posterior estimate conditional on the aggregate mean $\widehat{\mu}$, and Σ_{11} is the upper left element of the covariance matrix of the parameters, capturing the between-subject variance of μ_i. This model thus takes the form of a measurement error model, as often used in meta-analysis (see Brown, Imai, Vieider and Camerer, 2022, for the details of such a model). The posterior parameter will thus be shrunk towards the prior in proportion to the distance from the prior and its noisiness:

$$\mu_i = \frac{\Sigma_{11}}{\Sigma_{11} + \tau_i^2}\,\mu_i^{ML} + \frac{\tau_i^2}{\Sigma_{11} + \tau_i^2}\,\widehat{\mu}.$$

This has the advantage of discounting noisy outliers, which could otherwise be influential given the risk of overfitting when applying maximum likelihood techniques to small datasets (cfr. Bishop, 2006, chapter 3).[9] The method is thus explicitly geared at maximizing the *predictive*

[9] We here use the term 'maximum-likelihood estimation' to refer to individual-level estimations. Maximum likelihood techniques have also been adapted to estimate 'empirical Bayes' hierarchical models, which are however

performance of the estimates, even if this comes at the expense of individual-level fit, especially when the data are noisy. Indeed, it is easy to see that for noiseless observations, for which $\tau_i \to 0$, the posterior estimate $\mu_i \to \mu_i^{ML}$.

We estimate the model in Stan (Carpenter, Gelman, Hoffman, Lee, Goodrich, Betancourt, Brubaker, Guo, Li and Riddell, 2017) using Hamiltonian Monte Carlo simulations. We launch Stan from an algorithm R (https://www.rproject.org) using CmdStanR (Gabry and Češnovar, 2021). Covergence was carefully checked using best practices recommended by the Stan community. Online Appendix A provides the annotated code for the estimations we use and provides further technical details. Vieider (2024) provides a tutorial on the Bayesian estimation of decision models in Stan.

3 Results

3.1 Descriptive Statistics

We start by describing subjects' probabilistic beliefs about reciprocation. Before we start with the parametric estimation, however, we can try and get a feel for the data by examining choice proportions. A choice of special interest is the one between a bet on there being less than 50 reciprocators, or 50 or more. In this task, 58.7% of all subjects chose the option paying the prize if less than 50 subjects reciprocate, thus indicating a revealed majority belief that less than 50% of trustees will split the pie evenly. Nevertheless, only 26.8% of respondents reveal a belief that less than 25% of reciprocating trustees is more likely than there being between 25% and 50%. 60.1% believe that it is more likely that the number of reciprocators falls between 40 and 50 than between 50 and 60, and 65.4% believe that there being between 50% and 60% of reciprocators is more likely than there being between 60% and 70%. Table 3 in Appendix B.1 presents a complete list of the comparisons including the choice proportions.

Several choices are of the type $E_A = [0.5 - \frac{s}{2}, 0.5 + \frac{s}{2}]$ and $E_B = E_A^c$ (E_B is the complementary event of E_A). Figure 4 plots the proportion of choices of E_A as a function of s, i.e. as a function of the size of event E_A. It conveys two key results. First, on average subjects satisfy monotonicity with respect to s—the larger the event, the higher the choice proportion of that event. Second, when $s = 0.5$, the percentage of preference for E_A is significantly larger than 50%, meaning that the event $[0.25, 0.75]$ (there being between 25 and 75 reciprocators inclusive out of 100) is

fundamentally Bayesian. In that case, the intuition will be the same as presented here, except that the hyperpriors will be uniform over the entire support of the parameters, and that a interpretation of the estimated parameters as uncertain quantities is not warranted.

perceived as more likely than the complementary event $[0.25, 0.75]^c$. This shows that on average subject tend to have beliefs that peak towards the middle of the event space.

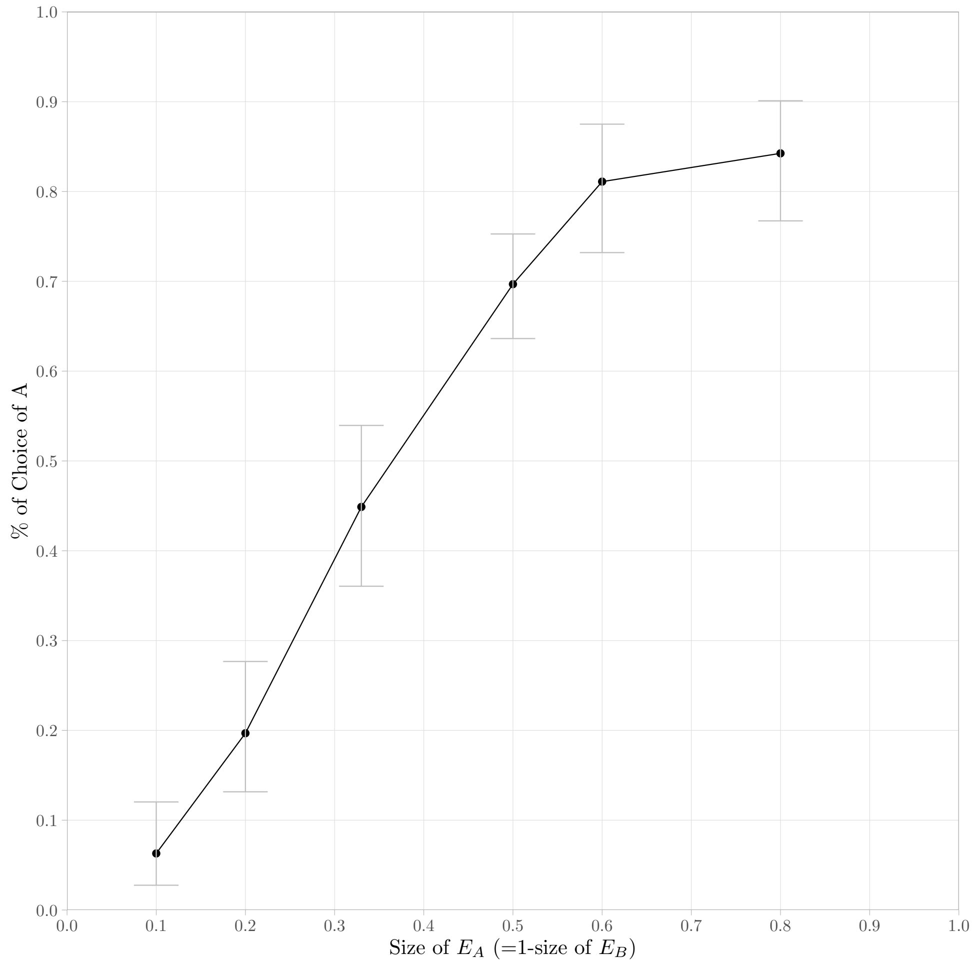

Figure 4: Choice between complementary events: $E_A = [0.5 - \frac{size}{2}, 0.5 + \frac{size}{2}]$ vs $E_B = E_A^c$

3.2 Estimation of belief distributions

Figure 5 shows a scatter plot of the mean belief about the proportion of reciprocating subjects, $\mu_i \triangleq \frac{\alpha_i}{\alpha_i+\beta_i}$, against the concentration, $\kappa_i \triangleq \alpha_i+\beta_i$, which we interpret as a measure of confidence in the mean belief. Panel A shows a scatter plot between the two measures, which can be seen to be orthogonal to each other ($\rho = -0.091$, $p = 0.306$, Spearman rank correlation). Few subjects believe that more than 75% of reciprocators split the pie evenly, whereas a sizeable minority believes that fewer than 25% do so. The median belief estimate is about 45%, and does not fall far from the true proportion of reciprocators, which is 0.496 (dashed blue vertical line). Subjects who themselves split the pie unevenly in their own favour tend to be more pessimistic. To the extent that they are optimistic, however, they seem to have low confidence in such beliefs, as witnessed by the large number of unequal splitters in the lower right part of the figure.

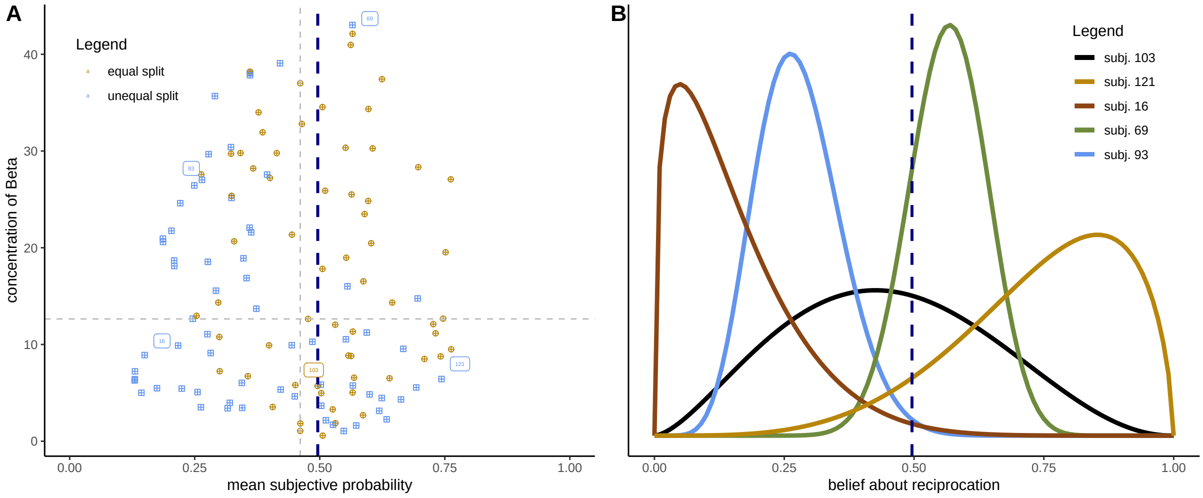

Figure 5: Belief distributions

Panel B shows some typical belief distributions. Subject 103 has a belief distribution that peaks just below 0.5, but which is extremely flat, showing very low confidence in that estimate. Subjects 93 and 69 have highly peaked distributions, but with very different means. Subject 16 is very pessimistic, albeit with a fair degree of uncertainty about that estimate, whereas subject 121 is very optimistic, but even less confident about that measure. Nevertheless, these two subjects are clearly separated in that the great majority of the subjective probability mass falls either above or below 0.5.

Table 1: Correlations between different belief measures

dep var: TE	WVS	trust2	trust3	quant.	mean Beta
WVS	1	–	–	–	–
helpful	0.392***	1	–	–	–
fair	0.386***	0.411***	1	–	–
quantitative	0.380***	0.306***	0.342***	1	–
mean Beta	0.170*	0.189**	0.180**	0.644***	1

The coefficient listed in the table indicate Spearman rank correlation coefficients. Stars indicate the following significance levels: * 10%; ** 5%; *** 1%.

Next, we explore the correlations between the different trust measures. Table 3 shows Spearman rank correlations between the different measures. The qualitative survey measures show correlation coefficients around 0.4 with each other, all of which are highly significant. Correlations of the survey measures with the introspective quantitative measure are somewhat lower and range between 0.3 and 0.4, but are still highly significant. The survey measures, however, correlate decidedly less well with the estimated mean belief, with correlations invariably below 0.2. By far the largest correlation, however, is the one observed between the estimated mean belief measure and the quantitative hypothetical measure, which comes to book at about 0.64. Revealed beliefs thus seem to line up with introspective, declared beliefs.

Beyond the mean estimates, we will use measures of the confidence subjects have in their own mean beliefs. Given the Bayesian setup we use, we can further describe the estimated parameters as probability distributions capturing posterior uncertainty about the estimated parameter values. This is paramount in the current setting, given the minimal set of stimuli that we used to identify beliefs.

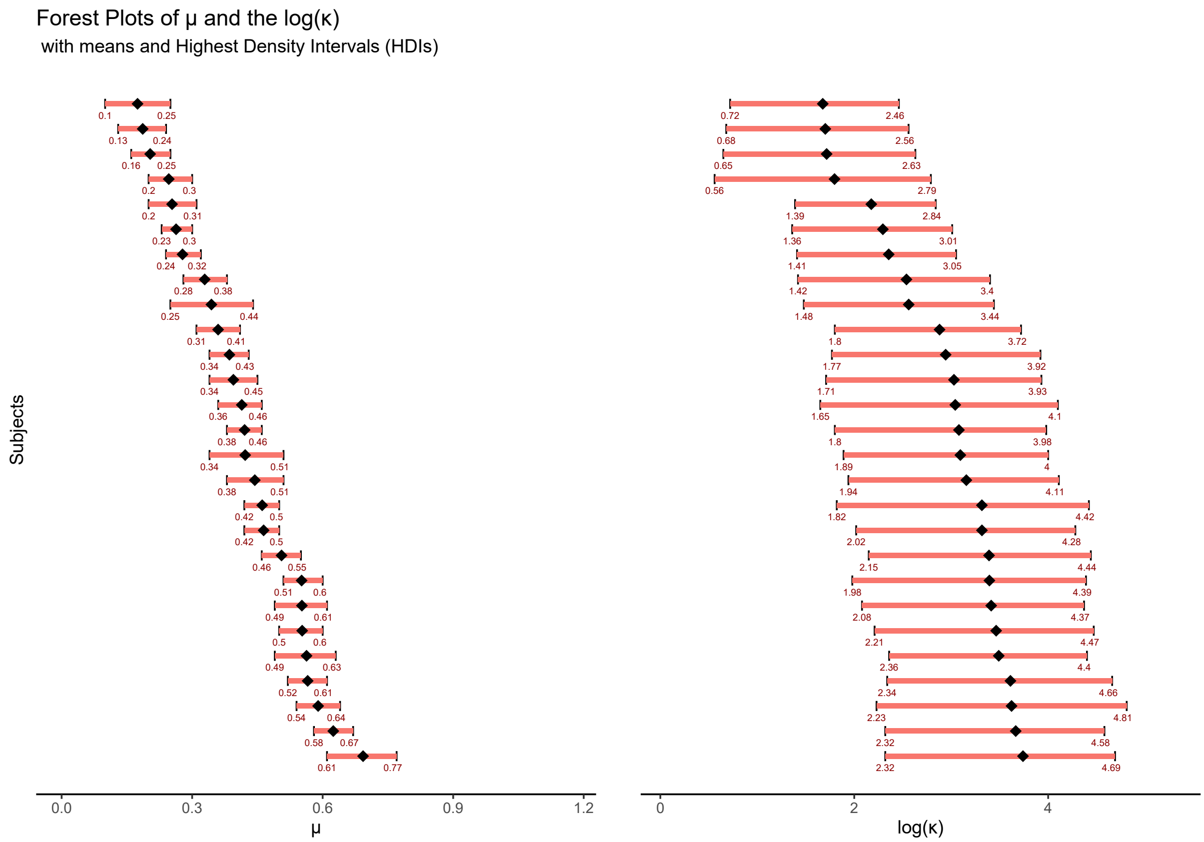

Figure 6: Forest Plots

Figure 6 plots the means and 95% highest density intervals (HDI)[10] of the posterior distributions of the mean, μ_i and logged concentration parameters, $log(\kappa_i)$, for 27 randomly selected subjects (the parameters of all subjects are reported in the Online Appendix). Posterior distributions of μ_i are symmetric since the mean estimates generally correspond to midpoints of

[10]A HDI is the narrowest interval that contains the specified amount of probability mass.

HDIs. The mean estimates for the concentration parameters, however, fall in the lower half of their posterior distributions which also have longer right tails, which is why we modelled their distribution using log-normals.[11]

3.3 Explaining trust equivalents

The principal measure used to validate our various measures of beliefs is their performance in explaining trust equivalents (TEs). Figure 7 shows the distribution of the trust equivalents to be explained, quantified as the midpoint between the last amount x for which a subject has chosen to trust, and the first sure amount for which they have chosen not to trust. Most of the trust equivalents we observe fall between Dhs 80 and Dhs 120, reflecting fairly low levels of trust. An outlying group of subjects displays high levels of trust. A few subjects either always chose the sure amount (2/127) or always chose to trust (4/127). Note that such preferences are legitimate in the trust game, as they may reflect extreme aversion to or preference for inequality (see Nunnari and Pozzi, 2022, for a meta-analysis of envy parameters). In what follows, our main interest resides in the extent to which different measures of probabilistic beliefs can account for the observed trusting behaviour.

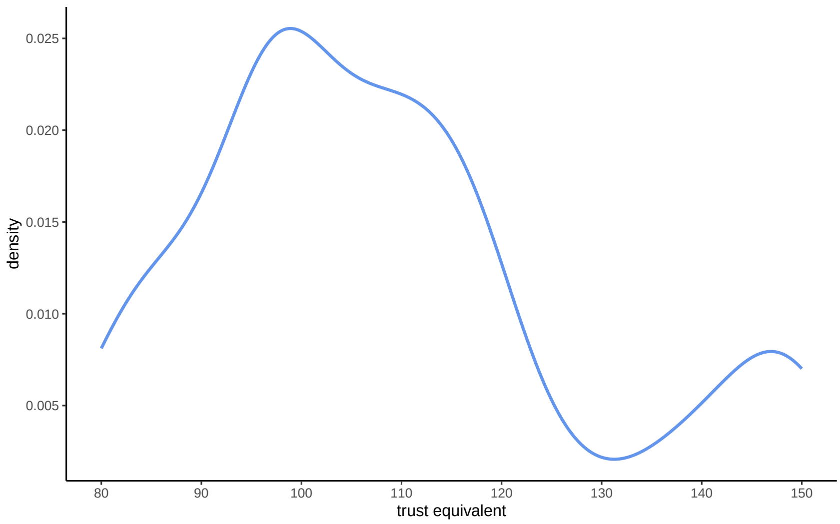

Figure 7: Belief distributions

In this section, we regress trust equivalents on various measures of beliefs and some control variables. We account for measurement uncertainty in our dependent variable by using an interval

[11]It is typical for variables constrained to be positive to follow skewed distributions. In our setting, using the log-normal ensures that this constraint is met. Limpert, Stahel and Abbt (2001) examined a variety of data sets across the sciences, and concluded that they could not find a single case in which the normal fit the data better than the log-normal.

regression. We furthermore use outlier-robust regressions throughout, taking the form of student-t distributions with 3 degrees of freedom (see Gelman, Carlin, Stern, Dunson, Vehtari and Rubin, 2014, section 17.5). This helps to guard against disproportionate effects of precisely estimated outliers (given that imprecise outliers will be discounted within our Bayesian hierarchical setting). The belief distributions based on choices between events are always estimated within one and the same econometric model containing the regression, so that the full posterior uncertainty about the belief parameters is taken into account in the regression.

We estimate the relationship between the parameters characterizing subjective belief distributions and ranges that contain the "true" values of trust in a *linear interval regression model.* To do this, let TE_i^* denote the "true" value of the trust equivalent for subject i. We assume that TE_i^* has mean $\mu_{TE_i^*}$, which is given by:

$$\mu_{TE_i^*} = \gamma_0 + \gamma_1\phi_i + \gamma_2\kappa_i. \tag{4}$$

The parameters of interest are thus γ_1 and γ_2, which are the coefficients of the mean and concentration parameters of belief distributions, respectively. Further, let l_i denote the highest sure amount that subject i gives up to choose to leave the decision to the trustee and h_i denote the lowest sure amount that subject i accepts instead of leaving the decision. The consecutive values l_i and h_i thus define the lower and upper bounds of the interval that contains the "true" but unknown value. In the case of subjects always choosing to trust or always choosing the sure amount, we leave the interval open-ended to one side. This setup therefore allows us to take the uncertainty in our dependent measure into account. Appendix A provides details and the provides the annotated code.

Table 2 shows the results of the TEs interval regressions, estimated by Bayesian procedures. Regressions I - III regress the TEs on qualitative survey questions about trust. Regression I uses the popular measure from the World Value Survey, although with an 11-point answer scale such as used e.g. in the European Social Survey. The question has significant explanatory power, with an R^2 of 0.056.[12] Regression II uses the question on whether “people try to be helpful”, and regression III the one suggesting that “people try to be fair” from the GSS.[13] These questions

[12] The measure of explained variance is calculated as $R^2 = 1 - \omega_m^2/\omega_0^2$, where ω_0^2 captures the variance in a model empty of covariates (intercept only), and ω_m^2 is the variance in the comparison model including various covariates.

[13] The precise formulation of the questions is: WVS: “Generally speaking, would you say that most people can be trusted or that you can’t be too careful in dealing with people?”; GSS1: “Would you say that most of the time, people try to be helpful, or that they are mostly just looking out for themselves?”; and GSS2: “Do you think that most people would try to take advantage of you if they got the chance or would they try to be fair?”

perform decidedly less well. They are neither statistically significant, nor do they explain the variance in TEs. Regression IV regresses trust on the answer to the quantitative introspective question of how many out of 10 participants are believed to choose an equal split on average. The measure is clearly predictive of trusting behaviour, and explains some 4.4% of variation in trust.

Table 2: Regression of trust on beliefs

dep var: TE	Reg. I	Reg. II	Reg. III	Reg. IV	Reg. V	Reg. VI	Reg. VII
WVS trust	1.612						
	(0.718)						
helpful		0.053					
		(0.676)					
try to be fair			-0.415				
			(0.597)				
equal split				1.659			
				(0.697)			
mean belief					23.725	24.816	24.864
					(8.978)	(8.795)	(9.062)
confidence in belief						0.385	0.437
						(0.182)	(0.175)
risk tolerance							1.753
							(0.554)
age, gender	NO	NO	NO	NO	NO	NO	YES
constant	99.061	104.449	106.486	96.967	93.751	87.634	64.718
	(2.888)	(3.350)	(2.972)	(3.514)	(4.349)	(4.955)	(17.738)
obs.	5588	5588	5588	5588	5588	5588	5588
subjects	127	127	127	127	127	127	127
R^2	0.056	0.002	0.011	0.044	0.077	0.208	0.252

Effects if age and gender are not shown for parsimony, and are not significant at conventional levels. However, gender = 1 is significantly negative at the 10% level.

Regression V uses the revealed preference measure for the mean belief about reciprocation, μ_i. This measure has a special place, inasmuch as it is the theoretically correct measure to use under subjective expected utility, as well as models that inherit its use of point-beliefs, such as prospect theory. The effect of the mean belief is highly significant, and explains 7.7% of the variance in TEs. Regression VI further adds the concentration of the Beta distribution as a measure of confidence in beliefs. While irrelevant based on SEU and PT, this measure is predicted to play an important role by multiple prior models. The effect of the mean belief measure results further reinforced by the insertion of the confidence measure. Furthermore, the confidence measure itself is a highly significant predictor of trust, which—ceteris paribus—increases in the confidence measure. Remarkably, inserting confidence into the regression brings the explained variance from 8% to over 20%, thus showing just how important it is to account for confidence in beliefs. Regression VII further adds the measures of declared risk tolerance validated by Dohmen, Falk, Huffman, Sunde, Schupp and Wagner (2011) and Vieider, Lefebvre, Bouchouicha, Chmura, Hakimov, Krawczyk and Martinsson (2015), and inserts additional controls for age and gender. Risk tolerance has a significant positive influence on trust, and brings the explained variance to 25%.

4 Discussion and Conclusion

Beliefs play a central role in economic decisions, yet the comparative performance of different belief measurements remains woefully under-investigated. Running a horse-race between introspective measures and incentivized measures of varying degree of complexity, Trautmann and van de Kuilen (2015) concluded that there is little if any evidence that more complex measures perform better. Here, we have proposed a simple, incentive-compatible measurement method for beliefs that relies on binary choices between bets on different events. Comparing the obtained measures to qualitative and quantitative introspective measures, we have shown that our method clearly outperforms those methods—certainly when a measure about confidence in beliefs is added. Yet our method remains very simple for subjects to understand.

The price to pay for the behavioural simplicity of our method lies in the rather more complex econometrics, which become necessary because we need to fit functional forms to belief distributions. In our view, this added complexity for the experimenter is a price well-worth paying for the simplicity in the elicitation design and for the incentive-compatibility of the tasks. The method we have proposed is portable across different elicitation environments, and can furthermore easily be adapted to different sorts of distributions if called for by the particular decision environment. Our findings have furthermore shown the importance of taking uncertainty seriously, both to rein in potentially noisy outliers, and to properly deal with posterior uncertainty about the estimated parameters.

Our findings also have important theoretical implications. In particular, the importance of confidence in beliefs can be taken to support models from the multiple-prior family, which postulate that aversion to ambiguity or uncertainty—two terms we use interchangeably in this context—will be driven by lack of confidence in average beliefs. While the evidence seems less compatible with models from the prospect theory family, our experiments were not devised as an explicit test of such a contrast. Indeed, the contrast between the two families of models may have been exaggerated, inasmuch as some models can be shown to converge to similar predictions based on particular functional forms (Ghirardato, Maccheroni and Marinacci, 2004). A careful theoretical investigation of these issues will have to wait for future work.

References

Abdellaoui, Mohammed, Aurélien Baillon, Lætitia Placido, and Peter P. Wakker (2011) 'The Rich Domain of Uncertainty: Source Functions and Their Experimental Implementation.' *American Economic Review* 101, 695–723

Abdellaoui, Mohammed, Han Bleichrodt, and Cedric Gutierrez (2024) 'Unpacking overconfident behavior when betting on oneself.' *Management Science*

Abdellaoui, Mohammed, Han Bleichrodt, Emmanuel Kemel, and Olivier L'Haridon (2021) 'Measuring beliefs under ambiguity.' *Operations Research* 69(2), 599–612

Aydogan, Ilke, Aurélien Baillon, Emmanuel Kemel, and Chen Li (2022) 'How much do we learn? measuring symmetric and asymmetric deviations from bayesian updating through choices'

Baillon, Aurélien (2008) 'Eliciting Subjective Probabilities through Exchangeable Events: An Advantage and a Limitation.' *Decision Analysis* 5(2), 76–87

Baillon, Aurélien, Han Bleichrodt, Umut Keskin, Olivier L'Haridon, and Chen Li (2018a) 'The effect of learning on ambiguity attitudes.' *Management Science* 64(5), 2181–2198

Baillon, Aurélien, Zhenxing Huang, Asli Selim, and Peter P Wakker (2018b) 'Measuring ambiguity attitudes for all (natural) events.' *Econometrica* 85, 1839–1858

Bellemare, Charles, Sabine Kröger, and Arthur Van Soest (2008) 'Measuring Inequity Aversion in a Heterogeneous Population Using Experimental Decisions and Subjective Probabilities.' *Econometrica* 76(4), 815–839

Bishop, Christopher M (2006) *Pattern recognition and machine learning,* vol. 4 (Springer)

Bohnet, Iris, and Richard Zeckhauser (2004) 'Trust , risk and betrayal.' *Journal of Economic Behavior & Organization* 55, 467–484

Bouchouicha, Ranoua, Jilong Wu, and Ferdinand M. Vieider (2023) 'Choice lists and 'standard patterns' of risk taking.' Technical Report, Ghent University Discussion Papers

Brown, Alexander L., Taisuke Imai, Ferdinand M. Vieider, and Colin F. Camerer (2022) 'Meta-analysis of empirical estimates of loss aversion.' *Journal of Economic Literature, forthcoming*

Carpenter, Bob, Andrew Gelman, Matthew D Hoffman, Daniel Lee, Ben Goodrich, Michael

Betancourt, Marcus Brubaker, Jiqiang Guo, Peter Li, and Allen Riddell (2017) 'Stan: A probabilistic programming language.' *Journal of Statistical Software* 76(1), 1–32

Danz, David, Lise Vesterlund, and Alistair J Wilson (2022) 'Belief elicitation and behavioral incentive compatibility.' *American Economic Review* 112(9), 2851–2883

Di Falco, Salvatore, and Ferdinand M. Vieider (2022) 'Environmental adaptation of risk preferences.' *Economic Journal* 132(648), 2737–2766

Dohmen, Thomas, Armin Falk, David Huffman, Uwe Sunde, Jürgen Schupp, and Gert G. Wagner (2011) 'Individual Risk Attitudes: Measurement, Determinants, and Behavioral Consequences.' *Journal of the European Economic Association* 9(3), 522–550

Engelmann, Dirk, Jana Friedrichsen, Roel van Veldhuizen, Pauline Vorjohann, and Joachim Winter (2023) 'Decomposing trust.' *Rationality and Competition Working Paper*

Fehr, Ernst (2009) 'On the Economics and Biology of Trust.' *Journal of the European Economic Association* 7(2-3), 235–266

Gabry, Jonah, and Rok Češnovar (2021) 'cmdstanr: R interface to'cmdstan'.' *URL: https://mc-stan. org/cmdstanr, https://discourse. mc-stan. org*

Gelman, Andrew, John B Carlin, Hal S Stern, David B Dunson, Aki Vehtari, and Donald B Rubin (2014) *Bayesian data analysis,* vol. 3 (CRC press Boca Raton, FL)

Ghirardato, Paolo, Fabio Maccheroni, and Massimo Marinacci (2004) 'Differentiating ambiguity and ambiguity attitude.' *Journal of Economic Theory* 118(2), 133–173

Glaeser, Edward L., David I. Laibson, Jose A. Scheinkman, and Christine L. Soutter (2000) 'Measuring Trust.' *The Quarterly Journal of Economics* 115(3), 811–846

Heinemann, Frank, Rosemarie Nagel, and Peter Ockenfels (2009) 'Measuring Strategic Uncertainty in Coordination Games.' *The Review of Economic Studies* 76(1), 181–221

Hossain, Tanjim, and Ryo Okui (2013) 'The binarized scoring rule.' *Review of Economic Studies* 80(3), 984–1001

Houser, Daniel, Daniel Schunk, and Joachim Winter (2010) 'Distinguishing trust from risk: An anatomy of the investment game.' *Journal of Economic Behavior & Organization* 74(1–2), 72–81

Karlan, Dean S. (2005) 'Using Experimental Economics to Measure Social Capital and Predict Financial Decisions.' *The American Economic Review* 95(5), 1688–1699

Klibanoff, Peter, Massimo Marinacci, and Sujoy Mukerji (2005) 'A Smooth Model of Decision Making under Ambiguity.' *Econometrica* 73, 1849–1892

Knack, Stephen, and Philip Keefer (1997) 'Does Social Capital Have an Economic Payoff? A Cross-Country Investigation.' *The Quarterly Journal of Economics* 112(4), 1251–1288

L'Haridon, Olivier, and Ferdinand M. Vieider (2019) 'All over the map: A worldwide comparison of risk preferences.' *Quantitative Economics* 10, 185–215

Li, Chen, Uyanga Turmunkh, and Peter P Wakker (2019) 'Trust as a decision under ambiguity.' *Experimental Economics* pp. 51–75

Li, Chen, Uyanga Turmunkh, and Peter P. Wakker (2020) 'Social and strategic ambiguity versus betrayal aversion.' *Games and Economic Behavior* pp. 272–287

Limpert, Eckhard, Werner A Stahel, and Markus Abbt (2001) 'Log-normal distributions across the sciences: Keys and clues.' *BioScience* 51(5), 341–352

Machina, Mark J, and David Schmeidler (1992) 'A more robust definition of subjective probability.' *Econometrica* pp. 745–780

Manski, Charles F. (2004) 'Measuring Expectations.' *Econometrica*

Nunnari, Salvatore, and Massimiliano Pozzi (2022) 'Meta-analysis of inequality aversion estimates'

Offerman, Theo, Joep Sonnemans, Gijs van de Kuilen, and Peter P. Wakker (2009) 'A Truth Serum for Non-Bayesians: Correcting Proper Scoring Rules for Risk Attitudes.' *The Review of Economic Studies* 76(4), 1461–1489

Rohner, Dominic, Mathias Thoenig, and Fabrizio Zilibotti (2013) 'Seeds of distrust: Conflict in uganda.' *Journal of Economic Growth* 18(3), 217–252

Schechter, Laura (2007) 'Traditional trust measurement and the risk confound: An experiment in rural Paraguay.' *Journal of Economic Behavior & Organization* 62(2), 272–292

Schotter, Andrew, and Isabel Trevino (2014) 'Belief elicitation in the laboratory.' *Annual Review of Econonomics* 6(1), 103–128

Tabellini, Guido (2010) 'Culture and Institutions: Economic Development in the Regions of Europe.' *Journal of the European Economic Association* 8, 677–716

Train, Kenneth (2009) *Discrete choice methods with simulation* (Cambridge; New York: Cambridge University Press)

Trautmann, Stefan T., and Gijs van de Kuilen (2015) 'Belief Elicitation: A Horse Race among Truth Serums.' *The Economic Journal* pp. 2116–2135

Trautmann, Stefan T, Gijs van de Kuilen, and Richard J Zeckhauser (2013) 'Social class and (un) ethical behavior: A framework, with evidence from a large population sample.' *Perspectives* *on Psychological Science* 8(5), 487–497

Vieider, Ferdinand M. (2024) ‘Bayesian estimation of decision models.’ Technical Report

Vieider, Ferdinand M., Mathieu Lefebvre, Ranoua Bouchouicha, Thorsten Chmura, Rustamdjan Hakimov, Michal Krawczyk, and Peter Martinsson (2015) ‘Common components of risk and uncertainty attitudes across contexts and domains: Evidence from 30 countries.’ *Journal of the European Economic Association* 13(3), 421–452

von Gaudecker, Hans-Martin, Arthur van Soest, and Erik Wengström (2011) ‘Heterogeneity in Risky Choice Behaviour in a Broad Population.’ *American Economic Review* 101(2), 664–694

ONLINE APPENDIX

A Estimation code and econometric details

Here, we explain our Stan code used to estimate our model. Vieider (2024) provides a general introduction to Bayesian estimation of decision making model, and a detailed tutorial how to estimate such models in Stan. Below, you can find the annotated Stan code. We launched this code from R using CmdStanR—see Vieider (2024) for detailed instructions and R code on how to do that. The example code below include the regression of trust equivalent on the measures of beliefs. Other controls can be added to the design matrix x, which should also include a column of 1s for the intercept. Comments are preceded by //

```
// designate and import data to be used in estimation:
data{
  int<lower=1> N; //number of obs.
  int<lower=1> Nid; // nr of unique identifiers (sequantial!)
  array[N] int id; // identifier (sequential integers!)
  array[N] int c; // dummy indicating choice of option A (c=1) or B (c=0)
  vector[N] pa; // lower bound of event A
  vector[N] qa; // upper bound of event A
  vector[N] pb; // lower bound of event B
  vector[N] qb; // upper bound of event B
  array[N] int cc;  // dummy indicating complementary event
  vector[Nid] trust_l; //lower bound of trust equivalent
  vector[Nid] trust_u; // upper bound of trust equiivalent
  int<lower=1> col; // nr. of columns in the design matrix
  matrix[Nid,col] x; // design matrix containing predictors
}

// declare main (primitve) parameters:
parameters{
  vector[2] means; // hierarchical means on trnasformed scale
  vector<lower=0>[2] tau; // vector of parameter SDs
  cholesky_factor_corr[2] L_omega; // Cholesky-decomposed covariance matrix
  array[Nid] vector[2] Z; // matrix of individual-level parameters, rescaled
  real<lower=0> sigma; // residual SD
  real bm; //coefficient on mean belief
  real bc; //coefficient on concentration
  real<lower=0> xi; // SD of robust regression
  vector[col] gamma; // vector of regression coefficiens
}

// transform parameters to use in model:
```

```
transformed parameters{
  matrix[2,2] Rho = L_omega * L_omega'; // obtain correlation matrix
  array[Nid] vector[2] pars; //matrix of individual-level parameters
  vector<lower=0,upper=1>[Nid] m; // individual-level means
  vector<lower=0>[Nid] k; // individual-level concentrations

  // loop to obtain parameters on original scale, to be used in model:
  for (n in 1:Nid){
    pars[n] = means + diag_pre_multiply(tau, L_omega) * Z[n];
    m[n] = inv_logit( pars[n,1] );
    k[n] = exp( pars[n,2] );
  }

// block estimating the model
}
model{
// auxiliary (local) vectors:
    vector[N] muA;
    vector[N] muB;
    vector[N] muB0;
    vector[Nid] mu;

    tau ~ exponential(5); //prior for parameters SDs
    L_omega ~ lkj_corr_cholesky(3); // prior for Cholesky-decomposed corr. matrix

    means[1] ~ normal(0, 10); // prior for mean aggregate belief (transformed scale!)
    means[2] ~ normal(0, 10); // prior for mean aggregate concentration (transformed scale!)

    sigma ~ normal(0, 20); // prior residual variance beliefs
    xi ~ normal(0, 200); // prior residual variance regression
    bm ~ normal(0, 200); // prior coefficient on mean belief
    bc ~ normal(0, 200); // prior coefficient on concentration
    gamma ~ normal(0, 200); // prior on regression coeff. vector

  for (n in 1:Nid)
      Z[n] ~ std_normal(); // prior on transformed individual-level pars

 // model estimation in loop:
    for ( i in 1:N ) {
    muB0[i] = beta_cdf(qb[i] | m[id[i]] * k[id[i]], // Beta CDF option B
                    (1 - m[id[i]]) * k[id[i]]) -
              beta_cdf(pb[i] | m[id[i]] * k[id[i]],
                    (1 - m[id[i]]) *k[id[i]]);
    muB[i] = muB0[i]^(1 - cc[i]) * (1 - muB0[i])^cc[i]; // Beta CDF complement of option B
    muA[i] = beta_cdf(qa[i] | m[id[i]] * k[id[i]], // Beta CDF option A
                          (1 - m[id[i]]) * k[id[i]]) -
```

```
            beta_cdf(pa[i] | m[id[i]] * k[id[i]],
                            (1 - m[id[i]]) * k[id[i]]);
    c[i] ~ bernoulli_logit((muA[i] - muB[i])/( sqrt(2) * sigma )) ; // model estimation
  }

// regression of trust equiivalents:
    for (n in 1:Nid){
    mu[n] = x[n] * gamma + bm * m[n] +  bc * k[n]; // regression

 // student-t robust regression:
    target += log_diff_exp(
    student_t_lcdf( trust_u[n] | 3 , mu[n], xi),
    student_t_lcdf( trust_l[n] | 3 , mu[n], xi));
    }
}
```

B Additional results

B.1 Binary Choice Tasks and Choice Proportions

B.2 Forest Plots of μ and κ for the remaining subjects

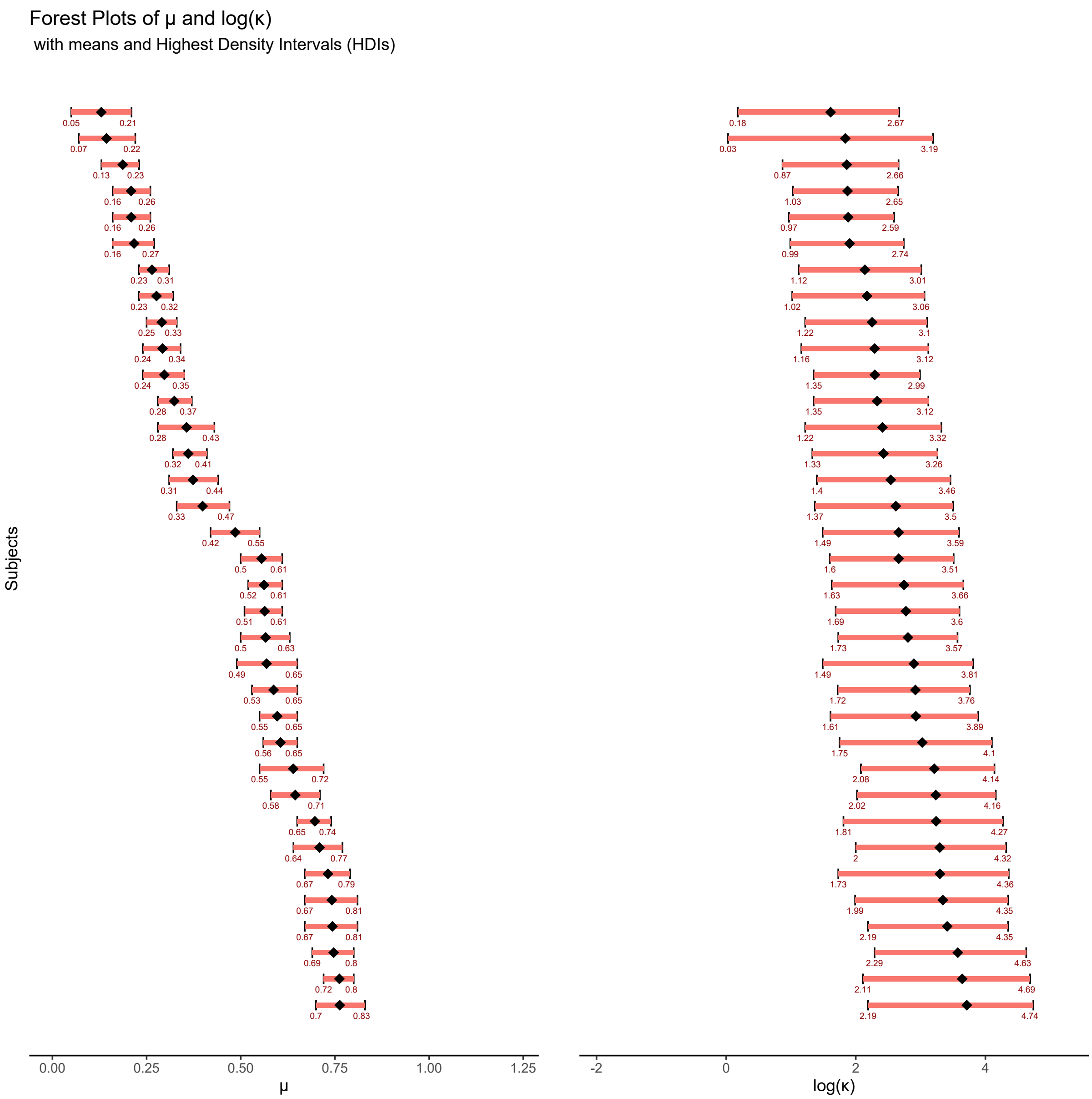

Figure 8: Forest Plots 1

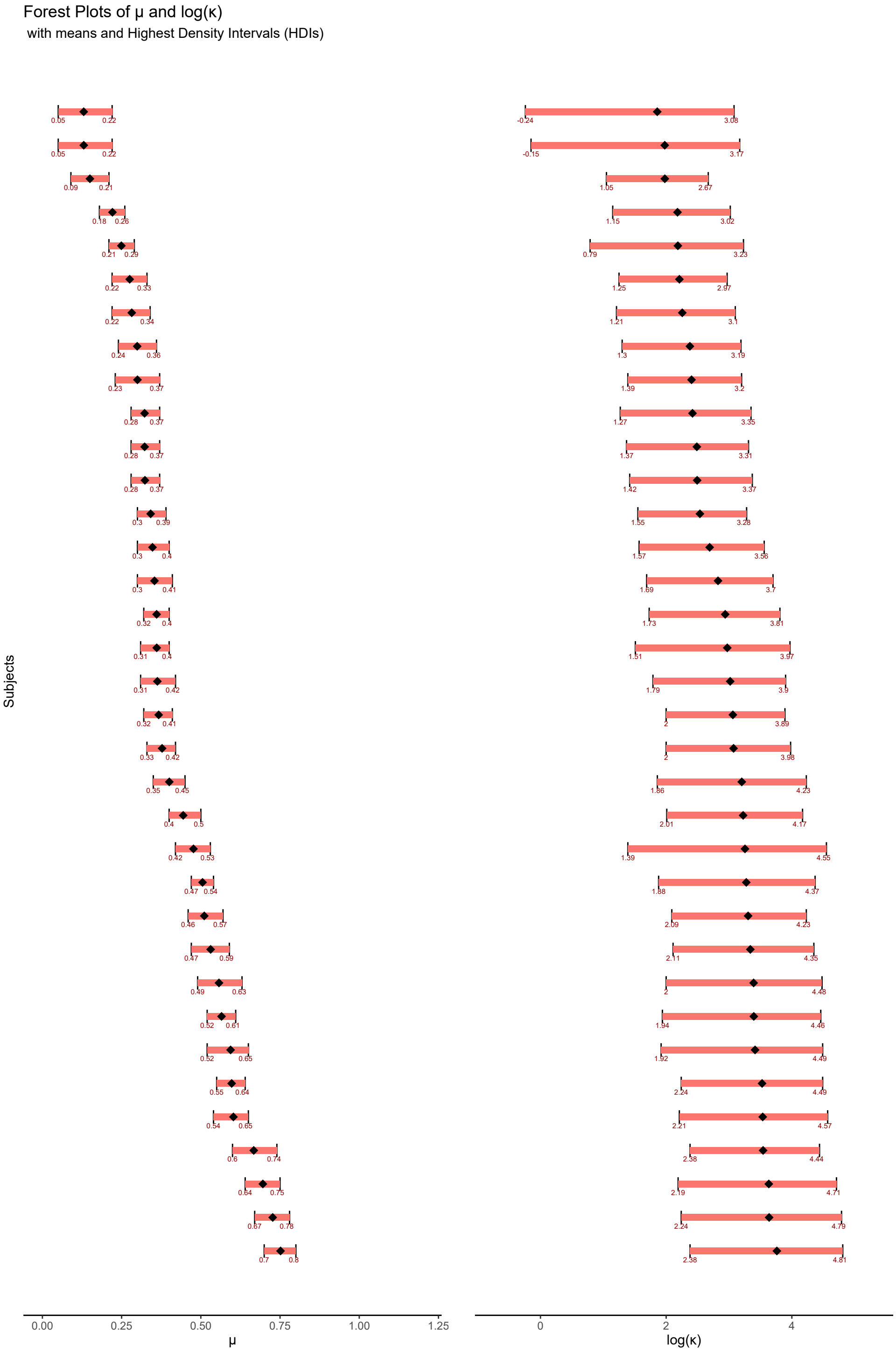

Figure 9: Forest Plots 2

Table 3: The List of Binary Choice Problems

Task ID	E_{A_j}	E_{B_j}	% Option A	Task ID	E_{A_j}	E_{B_j}	% Option A
1	[0,0.10)	(0.9,1]	60.6	21	[0.25,0.75]	[0,0.25)	80.3
2	[0,0.20)	(0.80,1]	56.7	22*	[0.25,0.75]	[0,0.25) ∪ (0.75,1]	69.7
3	[0,0.25)	[0.25,0.5]	26.8	23*	[0.25,0.75]	(0.75,1]	87.8
4	[0,0.25)	(0.50,0.75]	48.8	24	[0.30,0.40)	[0.40,0.50]	48.8
5*	[0,0.25)	(0.75,1]	59.8	25	[0.30,0.40)	[0.60,0.70]	58.2
6	[0,0.33)	(0.66,1]	62.9	26	[0.33,0.66]	[0,0.33)	61.4
7	[0,0.40)	(0.60,1]	62.2	27	[0.33,0.66]	[0,0.33) ∪ (0.66,1]	44.9
8	[0,0.45)	(0.55,1]	60.6	28	[0.33,0.66]	(0.66,1]	80.3
9*	[0,0.50)	[0.50,1]	58.6	29	[0.40,0.50]	[0.50,0.60]	60.6
10*	[0.10,0.30]	[0,0.10) ∪ (0.30,1]	21.2	30	[0.40,0.60]	[0,0.40)	42.5
11	[0.10,0.50]	[0,0.10) ∪ (0.50,1]	52.7	31	[0.40,0.60]	[0,0.40) ∪ (0.60,1]	19.7
12	[0.10,0.90]	[0,0.10)	85.8	32	[0.40,0.60]	(0.60,1]	68.5
13	[0.10,0.90]	[0,0.10) ∪ (0.90,1]	84.2	33	[0.45,0.55]	[0,0.45)	32.3
14	[0.10,0.90]	[0.90,1]	92.9	34	[0.45,0.55]	[0,0.45) ∪ (0.55,1]	6.3
15	[0.20,0.40]	[0,0.20) ∪ (0.40,1]	31.5	35	[0.45,0.55]	(0.55,1]	51.2
16	[0.20,0.80]	[0,0.20)	78.7	36	[0.50,0.60)	[0.60,0.70]	65.3
17	[0.20,0.80]	[0,0.20) ∪ (0.80,1]	81.1	37	[0.50,0.90]	[0,0.50) ∪ (0.90,1]	40.1
18	[0.20,0.80]	(0.80,1]	92.9	38	[0.60,0.80]	[0,0.60) ∪ (0.80,1]	18.9
19	[0.25,0.50)	[0.50,0.75]	60.6	39	[0.70,0.90]	[0,0.70) ∪ (0.90,1]	19.7
20	[0.25,0.50)	(0.75,1]	77.1				

Superscript * indicates that the task has been shown to subjects twice.

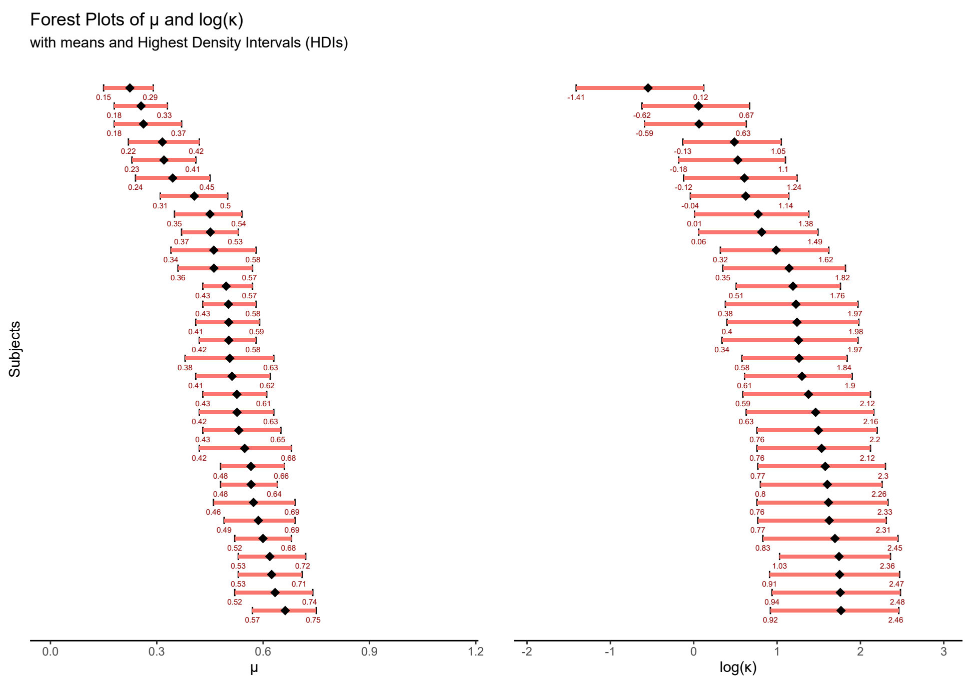

Figure 10: Forest Plots 3