

IZA DP No. 9856

Subjective Belief Distributions and the Characterization of Economic Literacy

Amalia Di Girolamo
Glenn W. Harrison
Morten I. Lau
J. Todd Swarthout

March 2016

Subjective Belief Distributions and the Characterization of Economic Literacy

Amalia Di Girolamo

University of Birmingham

Glenn W. Harrison

Georgia State University and IZA

Morten I. Lau

Copenhagen Business School and IZA

J. Todd Swarthout

Georgia State University

Discussion Paper No. 9856

March 2016

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0

Fax: +49-228-3894-180

E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

Subjective Belief Distributions and the Characterization of Economic Literacy*

We characterize the literacy of an individual in a domain by their elicited subjective belief distribution over the possible responses to a question posed in that domain. By eliciting the distribution, rather than just the answers to true/false or multiple choice questions, we can directly measure the confidence that an individual has about their knowledge of some fact. We consider literacy across several financial and economic domains. We find considerable demographic heterogeneity in the degree of literacy. We also measure the degree of consistency within a sample about their knowledge, even when that knowledge is imperfect.

JEL Classification: C91, D03, D83

Keywords: subjective beliefs, economic literacy

Corresponding author:

Morten I. Lau
Department of Economics
Copenhagen Business School
Porcelænshaven 16A, 1
DK-2000 Frederiksberg
Denmark
E-mail: mla.eco@cbs.dk

* We are grateful to the Danish Social Science Research Council (grant #12-130950) and the Center for Actuarial Excellence Research Fund of the Society of Actuaries for partial financial support.

Table of Contents

1. Procedures	-2-
A. Literacy	-2-
B. Demographics and Additional Measures	-6-
2. Belief Elicitation	-7-
3. The Measurement of Economic Literacy	-10-
A. Description of Results	-10-
B. Statistical Analysis of Results	-12-
4. The Consistency of Knowledge	-14-
5. Conclusions	-15-
References	-23-
Appendix A: Experimental Instructions (NOT FOR PUBLICATION)	-A1-
A. General Instructions	-A1-
B. Belief Elicitation Instructions	-A2-
C. Demographic and Other Hypothetical Questions	-A5-
Appendix B: Detailed Statistical Results (NOT FOR PUBLICATION)	-A8-

When we say that someone is literate we mean more than that they can just “read and write.” The term more generally indicates someone who is educated, whether by formal or informal means, and able to comprehend topics through words.¹ Characterizing and measuring the literacy of an individual requires then that we have some way of assessing *how* knowledgeable the person is about certain topics. There are some topics about which one can have “crisp” knowledge, in the sense of Boolean truth values. However, there are many domains of knowledge that one naturally expects varying levels of precision. We characterize literacy in terms of the subjective beliefs that someone has over possible responses to some question. By eliciting the subjective belief *distribution*, rather than just the answers to true/false or multiple choice questions, we can directly measure the confidence that an individual has about their knowledge of some fact.

Following Savage [1971][1972], we *define* subjective beliefs by the choices that individuals make when facing bets whose outcomes depend on those beliefs. The measurement of the literacy that someone has in a specific domain entails the elicitation of their subjective beliefs. For that task we conduct an experiment using proper scoring rules, which are simply structured bets offered to the individual by an observer (the experimenter). All of the elicited beliefs were incentivized and incentive-compatible, so that the subjects were making real choices with real economic consequences.

Our approach is to elicit the entire subjective belief *distribution* that an individual has, to ascertain how precise their knowledge is in response to some question. This extends and generalizes the prevailing approach to measuring literacy, which considers responses to (hypothetical) multiple choice questions (e.g., Lusardi and Mitchell [2007][2008][2013; §3]). For a specific question or domain, we are able to say “how literate” the person is, rather than just say that they are or are not literate. Of course, by asking a series of questions one can ascertain the fraction of correct answers for an individual with the traditional

¹ The *Oxford English Dictionary (Second Edition)* defines the adjective “literate” as someone who is “acquainted with letters or literature; educated, instructed, learned.” Remund [2010] offers a balanced account of the many definitions of literacy found in the academic and policy literature. Our focus on financial knowledge corresponds to the first of his categories of conceptual definitions of literacy (p. 279).

approach, but that requires one to pool responses over different questions which may span different knowledge domains.

The domains of interest to us are financial and economic knowledge. We consider a mixture of questions in which the correct answer involves the application of logical and grammatical principles, and questions in which the correct answer involves some specific fact. This reflects a trend in the measurement of literacy towards more than just the ability to draw logical or grammatical conclusions from information presented in the question itself, and to also consider awareness of facts that are of importance for the functioning of the individual.

A byproduct of our characterization is that we can also say something about the degree of common knowledge that a sample of individuals have about some proposition. Quite apart from whether or not a given individual knows the true answer with some precision, we often want to know if a group of individuals have the same degree of knowledge. In effect, we are able to operationalize several interpretations of what it means to have heterogeneous beliefs.

In section 1 we describe the experimental task we developed and employed with 120 subjects. In section 2 we review in detail the properties of the subjective belief elicitation procedure. In section 3 we present results on the degree of economic literacy of our subjects. In section 4 we consider the consistency of knowledge across subjects, and section 5 concludes.

1. Procedures

A. Literacy

We consider literacy in terms of four specific questions asked of each subject in an experiment. In each case there is a correct answer, and responses were elicited over a continuous range of possible answers presented in terms of 10 intervals or “bins.” A computer interface was used to present the belief elicitation tasks to subjects and record their choices, allowing them to allocate tokens in accordance with

their subjective beliefs. Figure 1 presents the interface.² The interface implements the Quadratic Scoring Rule discussed in section 2. Subjects could move the sliders at the bottom of the screen to re-allocate the 100 tokens as they wished, ending up with some distribution. The instructions explained that they could earn up to £20, as shown in Figure 2, but only by allocating all 100 tokens to one interval *and* that interval containing the true answer: if the true answer was just outside the selected interval, they would in that case receive £0. At the time of the experiments in December 2012, £20 was worth roughly \$32.

Subjects were rewarded for one of these belief elicitation tasks, with the task selected at random.³ The correct answer was revealed, and their earnings calculated according to the number of tokens allocated to the true interval in their elicited beliefs. For example, if the respondent had reported the beliefs in Figure 1, she would have been paid £16.25 if the correct answer was between 8% and 9.99%. As it happens, the correct answer here is 7.9%, so the subject would have actually received £11.25 since the correct answer was in the next lower interval, corresponding to unemployment rates between 6% and 7.99%.

The incentivized questions were as follows:

- **Q1: Interest Compounding.** “Suppose you had £100 in a savings account and the interest rate is 2% per year and you never withdraw money or interest payments. After 5 years, how much would you have on this account in total?” The correct answer is £110.40, and responses were elicited between £105 and £115 in intervals of £1.
- **Q2: Inflation.** “Suppose you had £200 in a saving account. The interest rate on your saving account was 1% per year and inflation was 2% per year. After 1 year, how much would be the value of the money on this account?” The correct answer is £197.96, and responses were elicited between £195 and £205 in intervals of £1. Due to a rounding error we rewarded subjects as if the correct answer was £198, and treat that as the “correct” answer.
- **Q3: Expected Lifetime for Men.** “Based on 2010 National Statistics, if a man lived to be 20 in the United Kingdom, how many more years would he expect to live? Note that this is not the age he would die at, but how many more years he would expect to live.” The correct answer is 59.1 years, and responses were elicited between 0 and 100 years in intervals of 10

² The instructions are reproduced in full in Appendix A. The interface was initialized with 10 tokens allocated to each bin.

³ The experiments also contained 4 belief elicitation questions that posed questions that tested knowledge of statistical principles. We do not evaluate those data here.

- years.⁴
- **Q4: Expected Lifetime for Women.** “Based on 2010 National Statistics, if a woman lived to be 20 in the United Kingdom, how many more years would she expect to live? Note that this is not the age she would die at, but how many more years she would expect to live.” The correct answer is 62.9 years, and responses were elicited between 0 and 100 years in intervals of 10 years.

The order of presentation of questions was randomized for each subject.

The first two questions are natural extensions of questions asked by Lusardi and Mitchell [2007][2008] in the *Health & Retirement Survey* (HRS) of 2004 in the United States.⁵ This survey is naturally representative of Americans over the age of 50. Our Q1 adapts the following question of theirs: “Suppose you had \$100 in a savings account and the interest rate was 2 percent per year. After 5 years, how much do you think you would have in the account if you left the money to grow: more than \$102, exactly \$102, less than \$102?” The main difference is that we ask for beliefs about the true answer over a wide range. Our Q2 adapts this question of theirs: “Imagine that the interest rate on your savings account was 1 percent per year and inflation was 2 percent per year. After 1 year, would you be able to buy more than, exactly the same as, or less than today with the money in this account?” Lusardi and Mitchell [2012; Table 2.1] report that only 67.1% and 75.2% of their sample gave the correct response to each question, respectively. These fractions drop significantly (their Figures 2.1a and 2.1b) as one considers Black and Hispanic respondents. When the same questions were posed to a nationally representative sample of young Americans, aged between 22 and 28 in Wave 11 of the *National Longitudinal Survey of Youth* conducted in 2007-2008, 79.3% and 54.0% gave the correct responses to the interest rate and inflation

⁴ At the time of the experiment we did not have access to the correct Life Tables, and instead subtracted 20 from the expected lifetime at birth from the UK Office of National Statistics to pay subjects that had this question selected. The difference for aggregates such as “all men” or “all women” is tiny, and did not affect the payments to any subject given that the bin intervals we used were in 10-year increments: there is a difference of 0.6 of a year for men (59.1 versus 58.5), and 0.5 of a year for women (62.9 versus 62.4).

⁵ A third question they asked was: *Do you think that the following statement is true or false? “Buying a single company stock usually provides a safer return than a stock mutual fund.”* This question was posed in order to understand if the individuals know how to diversify their investment. In a later Dutch national survey van Rooij, Lusardi and Alessie [2011] increased the set of questions posed to individuals. Apart from 5 questions aimed at characterizing “basic” financial literacy (p. 452), they added 11 questions to characterize “advanced” financial literacy (p. 454). Similar extensions were undertaken by Bateman, Eckert, Geweke, Louviere, Thorp and Satchell [2012] in surveys in Australia.

questions, respectively (Lusardi, Mitchell and Curto [2010; Table 1, p. 365]).⁶

The final two questions ask about a basic informational input to retirement planning: expected remaining lifetime, conditional on reaching the age of 20. Indeed, Smith, Taylor and Sloan [2001; p. 1126] call this “the most important subjective risk assessment a person can make,” although they were referring to own-mortality. We separate out the question for men and women, to ascertain if the differential expected mortality between the two is recognized by individuals. These questions do not condition on the health, income, or any other relevant characteristics of the individual that would affect expected mortality. One could easily extend these questions to elicit more precise beliefs about someone who is more similar to the subject.

The most widely used subjective beliefs about longevity come from the *Health and Retirement Survey*, which has asked a simple question since 1992 to respondents under the age of 65: “With 0 representing absolutely no chance, and 100 absolute certainty, what is the chance that you will live to be 75 years of age or older?” A comparable question asks the chance that they would live to be 85, and for respondents over 65 a variant asked the chances of them living 11-15 years more. In the 2006 wave of the *Health and Retirement Survey* a sub-sample was asked questions that elicited their beliefs about the population life tables: “Out of a group of [men/women] your age, how many do you think will survive to the age of X?” The value of X was 75 for those under 65 themselves, and 11-15 years older for those over 65. These questions are closer to those we asked, although we only conditioned on the single age 20.

Of course, these questions are not incentivized, and do not elicit information on the confidence of the subjective belief. However, Smith, Taylor and Sloan [2001] show that responses to this question are reasonably good predictors of future, actual mortality, even if they do not perfectly reflect new health information when updated. Perozek [2008] makes an even stronger case for the predictive value of these

⁶ Bateman, Eckert, Geweke, Louviere, Thorp and Satchell [2012] ask these questions of adult retirement savers in Australia, and find that 78.4% get the inflation question correct and 71.8% get the interest rate question correct.

subjective belief questions, arguing that responses to these questions actually outperform population life tables. In contrast, Elder [2013] stresses that only with the 2006 wave can one evaluate the actual predictions, as early respondents reach the target ages of 75 or 85. And in that respect he presents a sharply contrary view, arguing that the evidence supports a “flatness bias,” a “tendency for individuals to understate the likelihood of living to relatively young ages while overstating the likelihood of living to ages beyond 80.” He attributes this bias to a failure to recognize that mortality risk increases with age.

B. Demographics and Additional Measures

Apart from these incentivized subjective belief questions, which are our main focus, we asked subjects several hypothetical questions that have been widely used in the literature on cognitive abilities.

One is **Cognitive Reflection Test** (CRT) proposed by Frederick [2005], consisting of three questions:

- A bat and a ball cost £1.10 in total. The bat costs £1.00 more than the ball. How much does the ball cost?
- If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets?
- In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake?

The correct answers are £0.05, 5 minutes and 47 days. Many subjects responding to these hypothetical questions fail to “reflect” on some aspect of the information provided.⁷ The other hypothetical battery is known as the **Berlin Numeracy Test**, and is due to Cokely, Galesic, Schulz, Ghazal and Garcia-Retamero [2012]:

- Imagine we are throwing a five-sided die 50 times. On average, out of these 50 throws how many times would this five-sided die show an odd number (1, 3 or 5)?
- Out of 1,000 people in a small town 500 are members of a choir. Out of these 500 members in the choir 100 are men. Out of the 500 inhabitants that are not in the choir 300 are men. What is the probability that a randomly drawn man is a member of the choir? (Please indicate the

⁷ It is an interesting question whether one should care if subjects get the CRT questions right when they face incentives. In fact, many subjects do. But it could be argued that the CRT is designed to detect cognitive propensities to spot “minimally hidden” information in decision settings, and that this propensity is better detected when higher-order, executive brain functions are not engaged to earn money.

- probability in percentage).
- Imagine we are throwing a loaded die (6 sides) 70 times. The probability that the die shows a 6 is twice as high as the probability of each of the other numbers. On average, out of these 70 throws, how many times would the die show the number 6?
 - In a forest 20% of mushrooms are red, 50% brown and 30% white. A red mushroom is poisonous with a probability of 20%. A mushroom that is not red is poisonous with a probability of 5%. What is the probability that a poisonous mushroom in the forest is red?

The correct answers are 30, 25%, 20 and 50%.

In addition, a standard list of demographic questions were posed. These included age, sex, racial group, field of study, year of study, highest level of formal education expected to complete, current grade, citizenship, marital status, number of people in household, total income of the household, total income of parents, and smoking status.

2. Belief Elicitation

The decision maker in our experiment reports her subjective beliefs with a discrete version of a Quadratic Scoring Rule for continuous distributions, developed by Matheson and Winkler [1976].

Partition the domain into K intervals, and denote as r_k the report of the density in interval $k = 1, \dots, K$.

Assume for the moment that the decision maker is risk neutral, and that the full report consists of a series of reports for each interval, $\{ r_1, r_2, \dots, r_k, \dots, r_K \}$ such that $r_k \geq 0 \forall k$ and $\sum_{i=1..K} (r_i) = 1$.

If k is the interval in which the true value lies, then the payoff score is from Matheson and Winkler [1976; p.1088, equation (6)]:

$$S = (2 \times r_k) - \sum_{i=1..K} (r_i)^2$$

The reward in the score is a doubling of the report allocated to the true interval, and a penalty that depends on how these reports are distributed across the K intervals. The subject is rewarded for accuracy, but if that accuracy misses the true interval the punishment is severe. The punishment includes all possible reports, including the correct one.

Consider some examples, assuming $K = 4$. What if the subject has very tight subjective beliefs

and puts all of the tokens in the correct interval? Then the score is

$$S = (2 \times 1) - (1^2 + 0^2 + 0^2 + 0^2) = 2 - 1 = 1,$$

and this is positive. But if the subject has a tight subjective belief that is wrong, the score is

$$S = (2 \times 0) - (1^2 + 0^2 + 0^2 + 0^2) = 0 - 1 = -1,$$

and the score is negative. So we see that this score would have to include some additional “endowment” to ensure that the earnings are positive.⁸ Assuming that the subject has a very diffuse subjective belief and allocates 25% of the tokens to each interval, the score is less than 1:

$$S = (2 \times 1/4) - (1/4^2 + 1/4^2 + 1/4^2 + 1/4^2) = 1/2 - 1/4 = 1/4 < 1.$$

The tradeoff from the last case is that one can always ensure a score of 1/4, but there is an incentive to provide less diffuse reports, and that incentive is the possibility of a score of 1.

To ensure complete generality, and avoid any decision maker facing losses, allow some endowment, α , and scaling of the score, β . We then get the generalized scoring rule

$$\alpha + \beta [(2 \times r_k) - \sum_{i=1..K} (r_i)^2]$$

where we initially assumed $\alpha=0$ and $\beta=1$. We can assume different values of α and β to transform the payoffs to any alternative range of levels we may want.

In our experiment $K = 10$, and we do not know whether the subject is risk neutral. Indeed, the weight of evidence from past experiments clearly suggests that subjects will be modestly risk averse over the prizes they face. It is well-known that risk aversion can significantly affect inferences from applications of the Quadratic Scoring Rule when eliciting subjective *probabilities* over *binary* events (Winkler and Murphy [1970], Kadane and Winkler [1988]), and there are various methods for addressing these concerns.⁹ Harrison, Martínez-Correa, Swarthout and Ulm [2012] characterize the implications of the

⁸ This is a point of practical behavioral significance, but is not important for the immediate theoretical point.

⁹ For instance, see Köszegi and Rabin [2008], Holt and Smith [2009], Karni [2009] and Andersen, Fountain, Harrison and Rutström [2014].

general case of a risk averse agent when facing the QSR and reporting subjective *distributions* over *continuous* events, and find, remarkably, that these concerns do not apply with anything like the same force. For empirically plausible levels of risk aversion, one can reliably elicit the most important features of the latent subjective belief distribution without undertaking calibration for risk attitudes.

Specifically, they draw the following conclusions:

1. An individual reports having a positive probability for an event only if he has positive subjective probability for the event. So if the individual believes that unemployment is definitely below 12%, we would never see the individual reporting that it could be above 12%. Further, we can infer from Figure 1, for instance, that this subject truly attaches zero weight to the possibility of unemployment above 12%, no matter what his risk attitudes.
2. If an individual has the same subjective probability for two events, then the reported probabilities for the two events will also be the same if the individual is risk averse or risk neutral. So if the individual has a true, latent subjective probability of 0.1 that the unemployment rate is between 6% and 7.99%, and a true, latent subjective probability of 0.1 that it is between 10% and 11.99%, then the reported probabilities for these two intervals will be the same as well, as in Figure 1 (although typically not 0.1).
3. The converse is true for risk averse subjects, as well as for risk lovers. That is, if we observe two events receiving the same reported probability, we know that the true probabilities are also equal, although not necessarily the same as the reported probabilities.
4. If the individual has a *symmetric* subjective distribution, then the reported mean will be *exactly* the same as the true subjective mean, whether or not the subjective distribution is unimodal. Hence if we simply assume symmetry of the true distribution, a relatively weak assumption in some settings, we can elicit the mean belief directly from the average of the reported distribution.
5. The more risk averse an agent is, the more the reported distribution will resemble a uniform

distribution defined on the support of their true distribution. In effect, risk aversion causes the individual to report a “flattened” version of their true distribution, but never to report beliefs to which they assign zero subjective probability.

6. It is possible to derive the effect of increased risk aversion on the difference between the reported distribution and true distribution. Harrison, Martínez-Correa, Swarthout and Ulm [2012] show numerically that *a priori* plausible levels of risk aversion in laboratory settings implies no significant deviation between reported and true subjective beliefs in this setting.

Provided that our subjects exhibit the modest levels of risk aversion that are typically found in lab settings with similar stakes, these results provide the basis for using the reported distributions as if they are the true, subjective belief distributions.

3. The Measurement of Economic Literacy

A. Description of Results

In December 2012 we recruited 120 subjects from Durham University. The majority had major fields of study in Finance or some other Business area, and were completing a Master of Science degree. The average age was 24.4, 67% were women, and 85% were single and had never been married. Just over 73% were non-EU citizens, and 14% were current smokers.

Figure 3 provides a quick helicopter tour of the aggregate beliefs we elicited. More formal statistical tests are provided below. We observe very precise beliefs for the interest compounding question, which was relatively easy for our sample. Far less precision is observed for the other questions. Aggregate beliefs for the economic literacy questions tended to be unimodal, with most subjects having some sense of where the correct answer was, but with varying precision.

Figures 4 and 5 begin the evaluation of individual responses for the two financial literacy questions about interest compounding and inflation and the value of money. In each case we report the

correct answer, a “literacy index” and “concordance index,” the responses of three individuals selected to illustrate some differences in individual behavior, and the pooled distribution. We discuss the concordance index in the next section.

We construct a simple index of literacy, $L \in [0, 1]$, given by the fraction of 100 tokens that the individual allocates to the interval containing the true answer.¹⁰ This index does not need to be estimated: it is a direct transformation of the observed data. Thus we see a value of $L=1$ for subject #1 in Figure 5, the interest compounding domain, since this subject allocated all 100 tokens to the interval containing the correct answer. Many subjects did exactly the same thing in this case, but subject #10 and subject #11 show how a few hedged their bets, quite literally. For subject #10, 60 of the 100 tokens were in the interval containing the correct answer, so $L=0.60$ for this subject.

In Figure 5 we see that subject #1 has a literacy index of zero, since she allocated all 100 tokens to the interval just to the left of the correct answer for the inflation question. In this domain, subject #5 has a literacy index of 0.26 since 26 out of 100 tokens were allocated to the correct interval. By being less dogmatic, subject #5 exhibited greater literacy than subject #1. Of course, one might want to argue that subject #1 was very close to the correct answer, but countering that claim is the subject’s choice, implicitly saying that she was certain of her answer. If indeed she has some imprecision, that should have led her to report a non-degenerate distribution.

Figures 6 and 7 consider the other questions, about expected remaining lifetime for men and women. Figure 6 shows detailed responses for 11 individuals to the question about men, since we

¹⁰ There are many other literacy indices that could be constructed from these data. For instance, a natural metric would be the amount of *money* that the subject had “left on the table” by not allocating all 100 tokens to the correct answer. This reflects the QSR underlying the incentives provided to subjects to be more precise, and is therefore a non-linear transformation of all token allocations. More generally, there is one way to be completely correct about these questions, and 4,263,421,511,270 ways to be wrong. If there are t tokens and b bins, then there are $(t+b-1)!/t!(b-1)!$ possible allocations in each of our elicitation tasks. Only one of these is completely correct. If someone has a literacy index $L=0$ then there are still $(100+9-1)!/100!(9-1)! = 352,025,629,371$ ways to respond.

observe considerable heterogeneity in this domain compared to the financial literacy questions. The imprecision for subject #3, #5 and #7 is substantial, and leads one to speculate if it shows up in their savings behavior or retirement planning. Figure 7 shows the differences in the aggregate distribution between the questions about men and women, to gauge if the longer expected lifetime of women compared to men is generally understood. Indeed, we see that this increment is detected in aggregate, so we can say that we observe a reasonable “literacy of crowds” on this issue.

Finally, Figure 8 collates information on the distribution of the literacy indices L across the domains considered. The vertical, dashed line is the average of each distribution, for reference. The distribution for the interest compounding question, in the top left panel of Figure 8, is what one would normatively like to see: almost universal high-literacy. However, one can visually infer that this is the exception across these domains.

B. Statistical Analysis of Results

A natural statistical model for directly evaluating the beliefs data is interval regression. In this specification the dependent variable refers to the intervals given by our elicitation “bins.”¹¹ In all cases we control for sex, age, marital status, race, whether a Finance major, whether a non-EU citizen, whether a current smoker, and the scores on the CRT and Berlin Numeracy Test. The age and test scores are all normalized to have mean zero and unit standard deviation. Detailed estimates are provided in Appendix B.

The interval regression models for each belief question show the following statistically significant demographic effects:

¹¹ Interval regression allows one to identify copen intervals with $-\infty$ as a lower bound or $+\infty$ as an upper bound, but that is not generally appropriate for our elicitation tasks. We evaluate the effects of using such copen bounds for the interest compounding and inflation questions, and it makes no difference to our conclusions.

- Older individuals are slightly more literate in the interest compounding domain, although there is very little variation in the dependent variable here (Figure 4).
- Women have slightly lower literacy in the inflation domain.
- Whites are much more literate on the expected remaining life years for men (and also for women, although the effect for women is not statistically significant).
- Older individuals are substantially more literate on the expected remaining life years for women.

We evaluate the systematic effect of demographics more formally below by pooling across measures of literacy L in different domains.

To assess the effects of demographics on literacy we estimate an ordered logit model of the literacy index L , but collapse the index to three values for ease of interpretation. One value corresponds to $0 \leq L \leq 1/3$ and might be called “illiterate,” another value corresponds to $1/3 \leq L \leq 2/3$ and might be called “semi-literate,” and a final value corresponds to $L > 2/3$ and might be called “literate.” Table 1 reports the marginal effects¹² of the listed covariates on the probability of being illiterate, semi-literate and literate, as defined here.

These effects are normalized to the high literacy found in the interest compounding question, so it is not surprising to see most of the dummies for individual questions showing an increased effect on the probability of *illiteracy* by comparison. For instance, compared to the interest compounding question, the effect of asking the inflation question was to increase the probability of illiteracy ($L \leq 1/3$) by 0.10.

Focusing only on statistically significant effects, in this sample women are more likely to be illiterate and less likely to be literate than men. Older subjects are more likely to be semi-literate and literate. Marriage is associated with a striking increase in illiteracy. Being a Finance major is associated with heightened literacy. Being a non-EU citizen is associated with a much lower level of literacy compared to EU citizens. Finally, the Berlin Numeracy Test is associated with a significantly higher probability of being literate. These demographic effects are a mix of the expected (e.g., the Berlin

¹² These are the *average* marginal effects, evaluated by considering all of the actual values of the non-target variables and averaging the marginal effects across those values. The alternative is to consider the average values of the non-target variables and reporting the marginal effect at that single set of values.

Numeracy Test) and the unexpected (e.g., being single). The significance of these demographic effects also points to the heterogeneity of literacy across these domains.

4. The Consistency of Knowledge

Although literacy is a capacity that is naturally measured for the individual, it obviously impacts the extent to which knowledge about something is shared.¹³ If someone has a poor level of literacy in some domain, the natural question is whether that is consistent with the knowledge that others have. The immediate consequences for behavior when there are heterogeneous beliefs are by now well-studied, such as in models of asset pricing in finance (e.g., Shefrin [2008]), game theoretical interaction, and rational expectations.

These ideas are also familiar from linguistics. The process of learning a language involves the disambiguation of utterances (Allen [1995]). And many linguists discuss language use as intentionally constrained by norms of communicating understanding, which is to say greater literacy (Grice [1989], Clark [1992]). Hence one naturally seeks some measure of shared literacy. Is the uncertainty over some fact in a given domain shared, or is it a domain in which one can clearly identify “experts” and “novices?” We propose a simple measure that can allow us to address that question.

Any measuring instrument can be compared against another measuring instrument. Examples include weight scales, political opinion polls, or medical judgements about diagnoses. In our case we consider the subjective beliefs about some fact, and seek to measure their consistency. In the biostatistics

¹³ When measuring the literacy of a *household*, how does one account for the heterogeneity of levels of literacy within the household? The concept of effective literacy, developed by Basu and Foster [1998], considers this important dimension of what might also be called “social literacy.” Obviously the measurement issues goes beyond the household, and includes any social network used by individuals for making decisions. One approach focuses on potential literacy, and defines literacy in terms of the most literate person in the household. This adjustment to naïve measures of literacy at the individual level is easy to make, and provides a valuable upper bound. One could, however, similarly define a lower bound if the household power relationships lead to the least literate person imposing his or her will on the household decision in some setting. Using our tools for characterizing the production function for social literacy in these settings is an important extension.

literature a popular concordance index ρ_c has been developed by Lin [1989][2000]. It combines the familiar notion of correlation from a Pearson inter-class correlation coefficient with allowance for bias, and is virtually identical to measures of intra-class correlation (Nickerson [1997]). The concordance index is bounded between ± 1 , with the usual interpretation that $\rho_c = 1$ indicates perfect concordance, and smaller values indicate poorer concordance.

In Figure 4, for instance, we evaluate the concordance index for each subject with respect to the pooled belief distribution on the interest compounding question, and then also report the average value of the index over all 120 subjects. Even though subject #1 in that setting had a literacy index value of 1, since she gave the correct responses, her concordance with the group was slightly less than 1 (0.972) because some people in the group did not have perfect literacy in this domain (e.g., subjects #10 and #11, shown in Figure 4). Taking a less extreme case, such as the inflation and value of money question in Figure 5, we see much lower levels of concordance. Subject #5, even though less precise than subject #1 and subject #3, was more consistent with the beliefs that everyone else had.

Moving to the distribution of concordance indices in each domain, Figure 9 shows the heterogeneity of beliefs we elicited. The interest compounding question, in Figure 9, is again an outlier, showing considerable literacy in this sample (from Figure 8) and hence considerable consistency.

5. Conclusions

Literacy is a concept that is widely discussed, and clearly at the core of understanding economic behavior in modern societies. We propose a characterization of literacy using the familiar notion of a subjective belief distribution over questions in a certain domain.¹⁴ We demonstrate how these belief

¹⁴ The modern policy literature on literacy stresses the concept of “capability,” which is the extent to which individuals use their knowledge, as distinct from being able to answer abstract questions successfully. The concept of capability seeks to characterize if someone is able to function in a certain domain. This raises many subtle, interesting issues. First, it is not obvious that someone must know the right answer in order to be able to function in some domain robustly. There are some task domains where the payoffs are very “flat,”

distributions can be elicited in an operational, incentive-compatible manner from individuals. We show that there is considerable heterogeneity in literacy levels over economic and financial domains, and across observable demographics.

Immediate extensions of our approach are to consider broader samples and other domains of literacy, as well as the effects of controlled interventions on the distribution of literacy. Evaluation of the consequences of imperfect illiteracy can be undertaken by studying the choices made, or avoided, in related tasks that rely on literacy in that domain.¹⁵ Do semi-literate individuals avoid welfare-improving choice domains for fear that they might make serious mistakes? Straightforward extensions to consider social literacy are also important and natural using our characterization.

in the sense that large errors in the specific choice lead to virtually the same expected payoff as more refined choices. Second, it is also not obvious that someone must infer the right answer by applying grammatical or logical algorithms in order to make good choices: heuristics might do very well in many domains, whether or not there is a flat payoff at work in the region of choice. Third, the concept of capability raises the issue of domain-specific knowledge, which goes beyond the “reasoning from first principles and the information in front of you” approach that characterizes most analyses of literacy. Someone might be a wizard at applying Bayes Rule, but simply have an incorrect prior belief about some base rate. Such a person would typically be deemed statistically literate but not capable.

¹⁵ Another byproduct of our approach to literacy is the ability to formally characterize what it means for individuals in a sample to be “overconfident” about their literacy, in the sense that they overplace themselves relative to others. Moore and Healy [2008; p.508], Merkle and Weber [2011; p.264] and Benoit and Dubra [2011; p. 1605] explain why elicited subjective belief *distributions* are needed to evaluate such hypotheses. These characterizations would then have immediate application to the related “unskilled and unaware of it” hypothesis of Kruger and Dunning [1999] that those that tend to be less literate also tend to be the least aware of their relative disability.

Figure 1: Belief Elicitation Interface

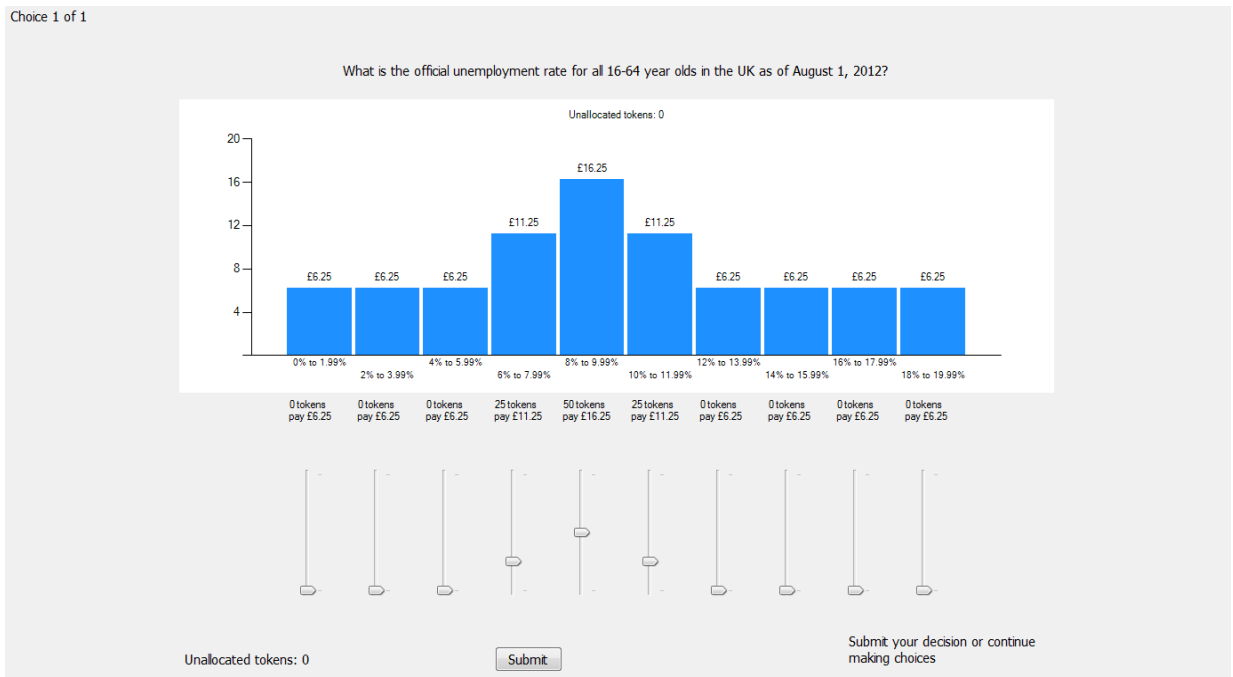


Figure 2: Possible Belief Elicitation Response

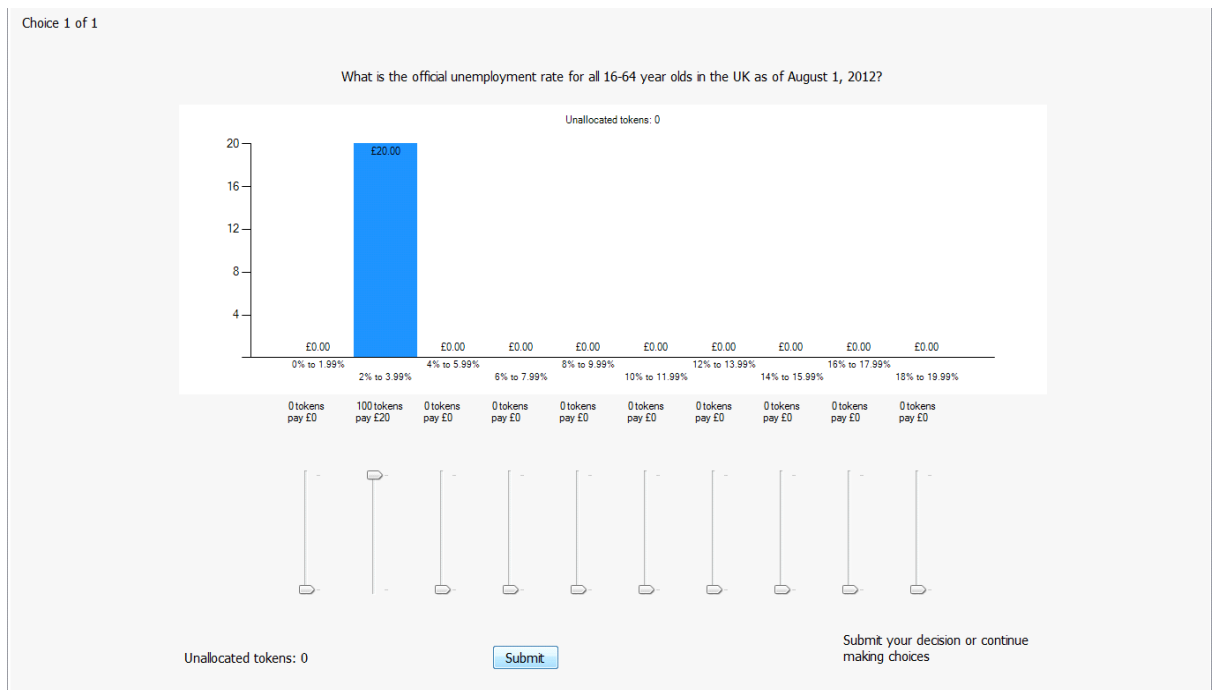


Figure 3: Pooled Subjective Beliefs

Labels in left (right) panel show lower bound (midpoint) of elicited interval

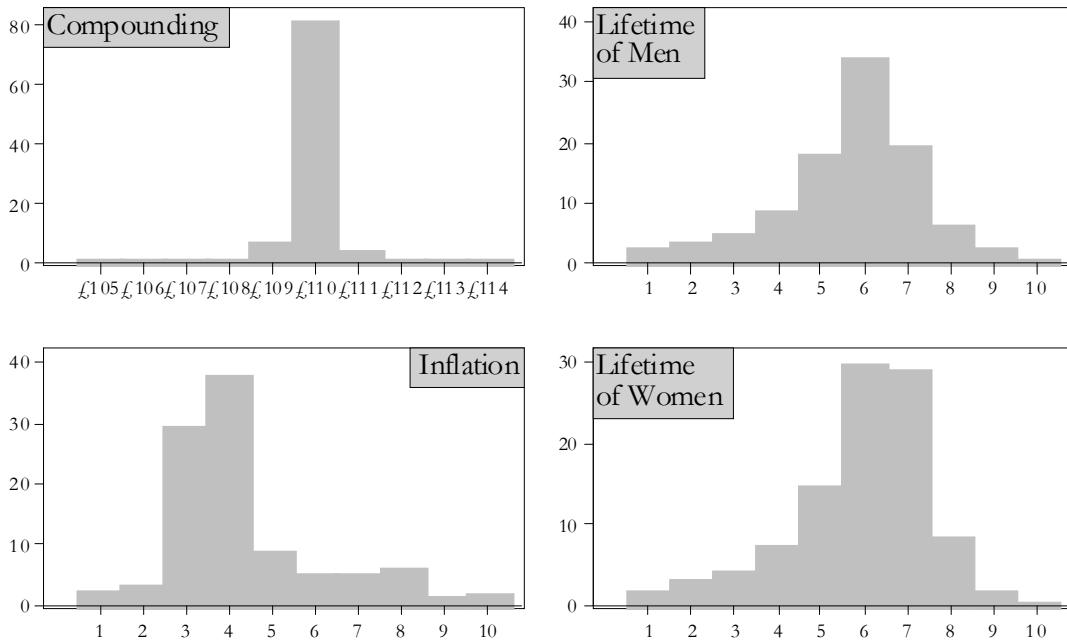


Figure 4: Subjective Beliefs of Three Subjects to Interest Rate Compounding Question

Correct Answer is £110.40

Average Literacy index $L = 0.81$ and average Concordance index $\rho_c = 0.879$

Labels show lower bound: e.g., £105 refers to the interval £105 to £105.99

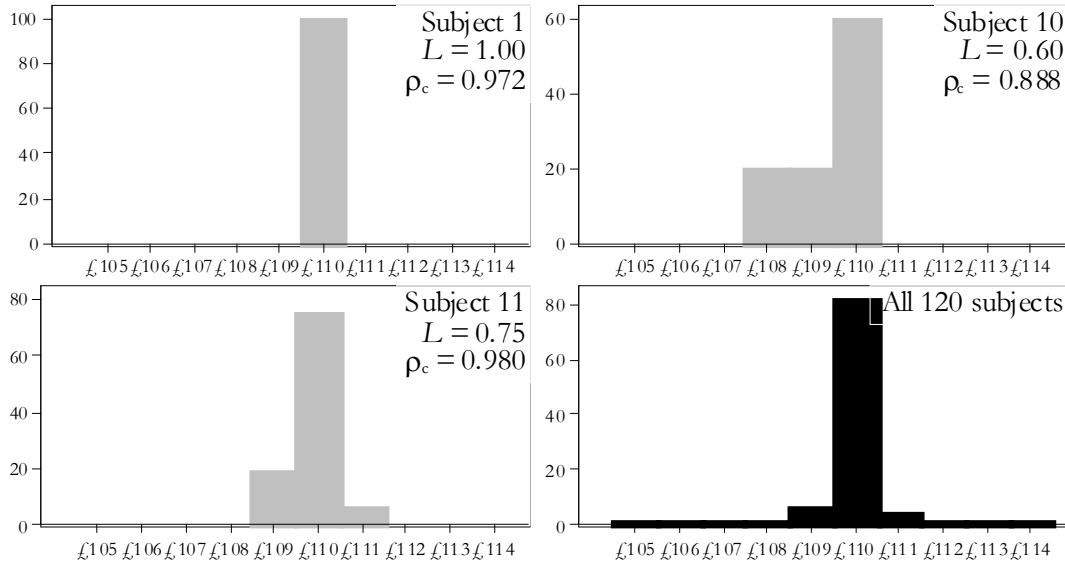


Figure 5: Subjective Beliefs of Three Subjects to Inflation and the Value of Money Question

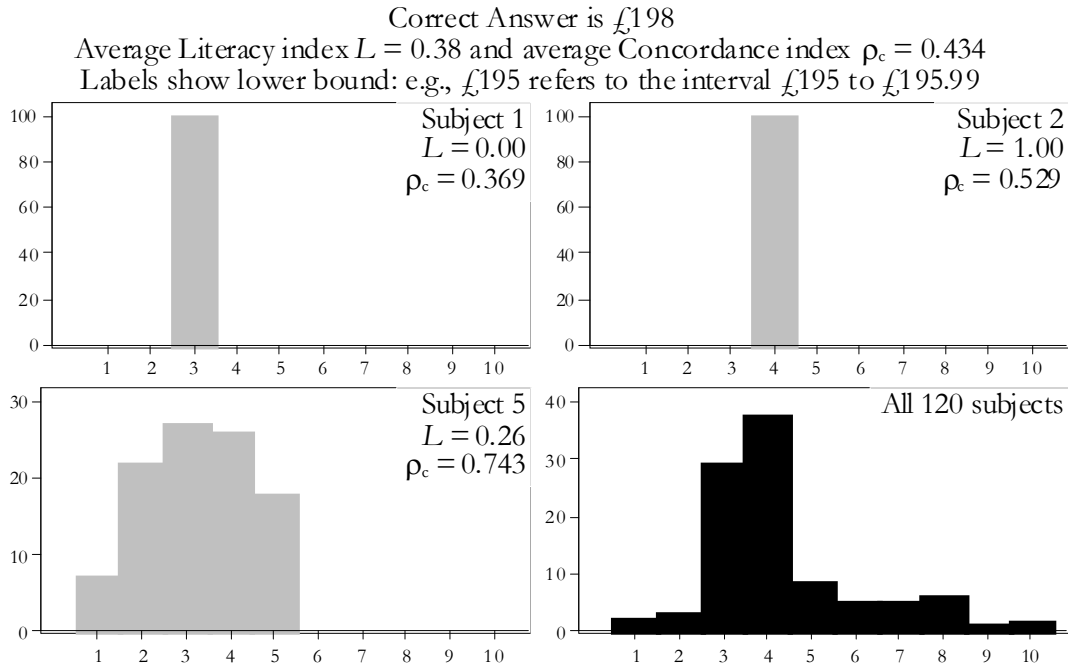


Figure 6: Beliefs on the Remaining Lifetime of Men

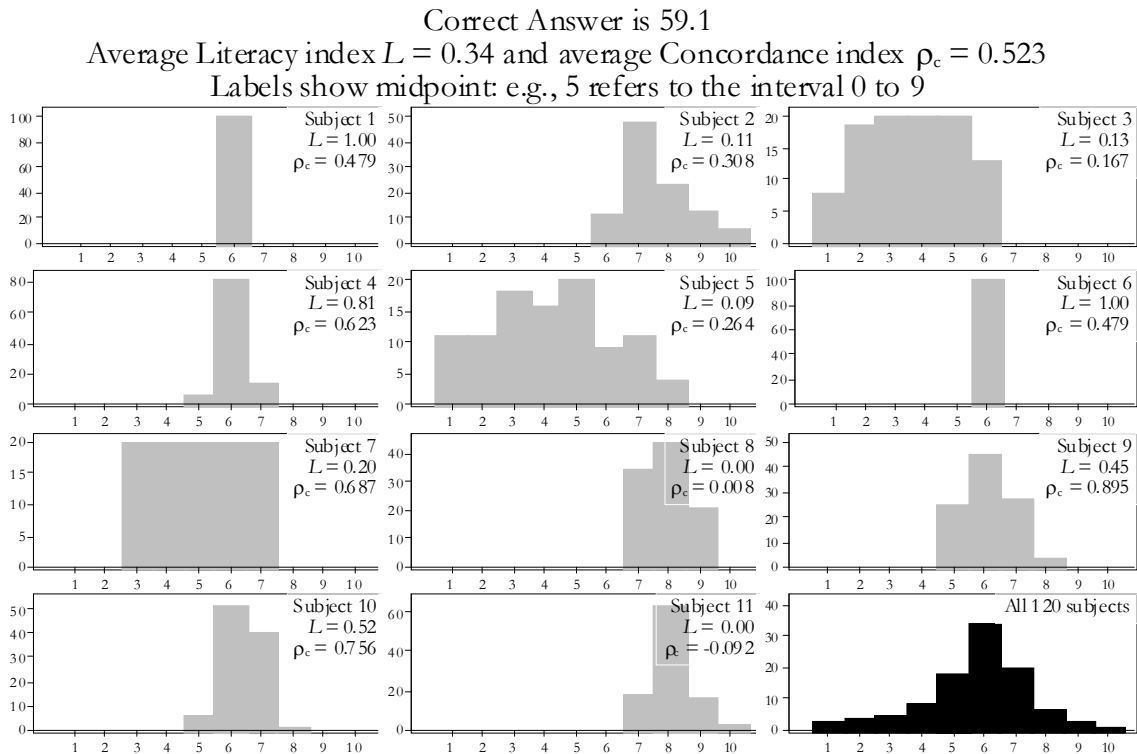


Figure 7: Elicited Beliefs For Remaining Lifetime

Correct answer is 59.1 years for men and 62.9 years for women
 Average Literacy index $L = .34$ for men and $.29$ for women

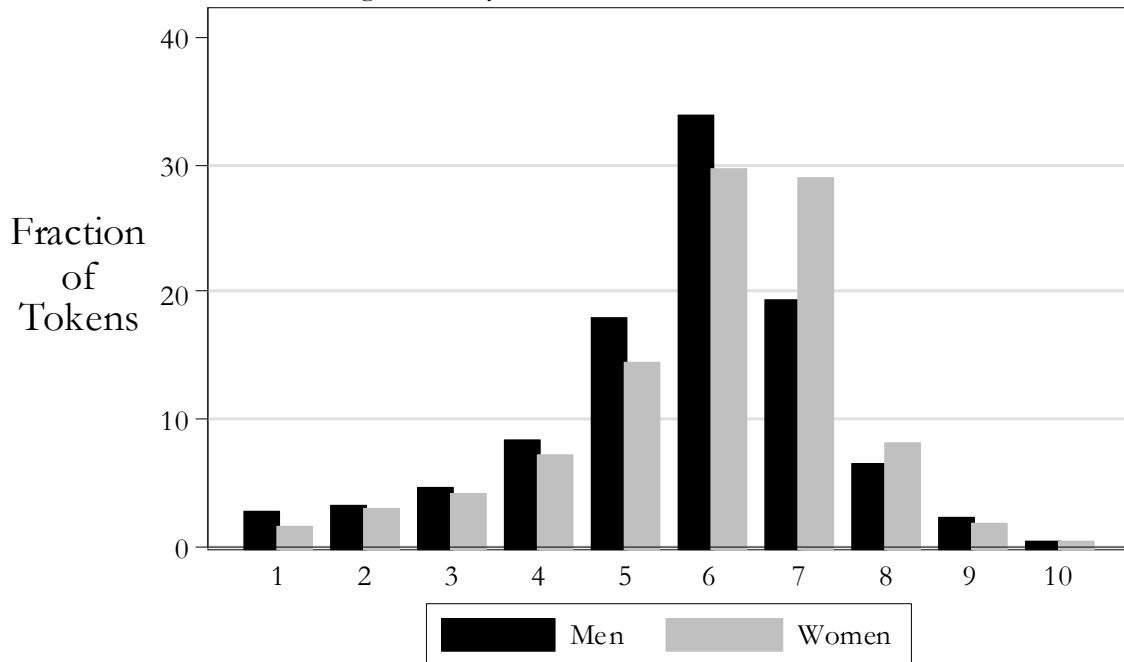


Figure 8: Literacy Indices

Dashed line is the average

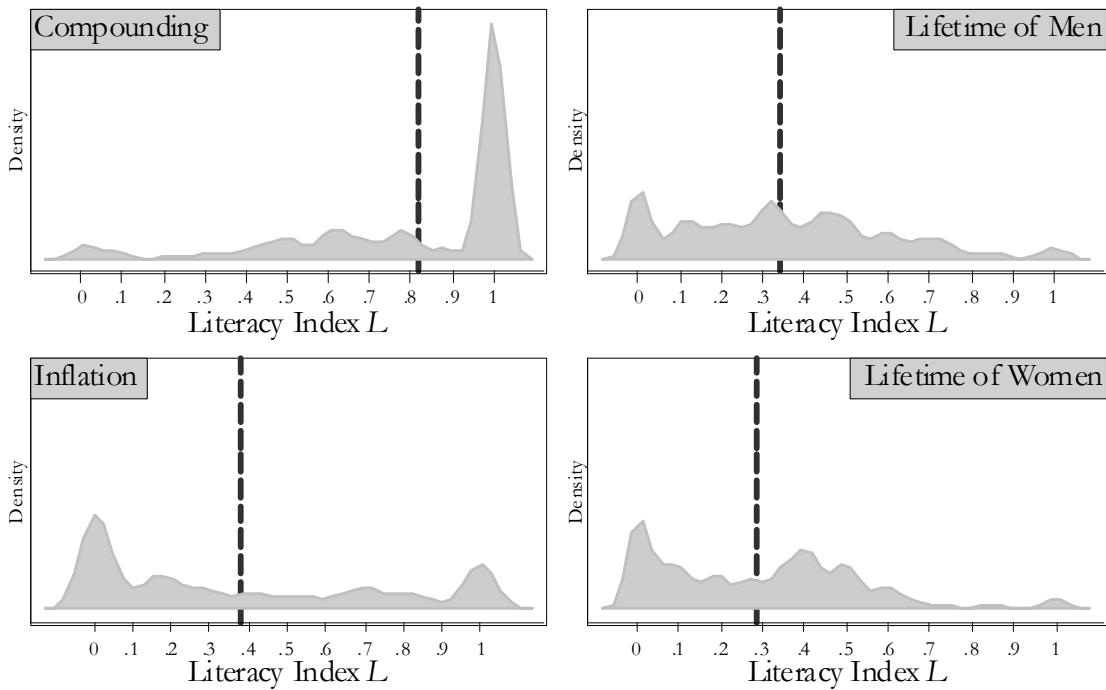
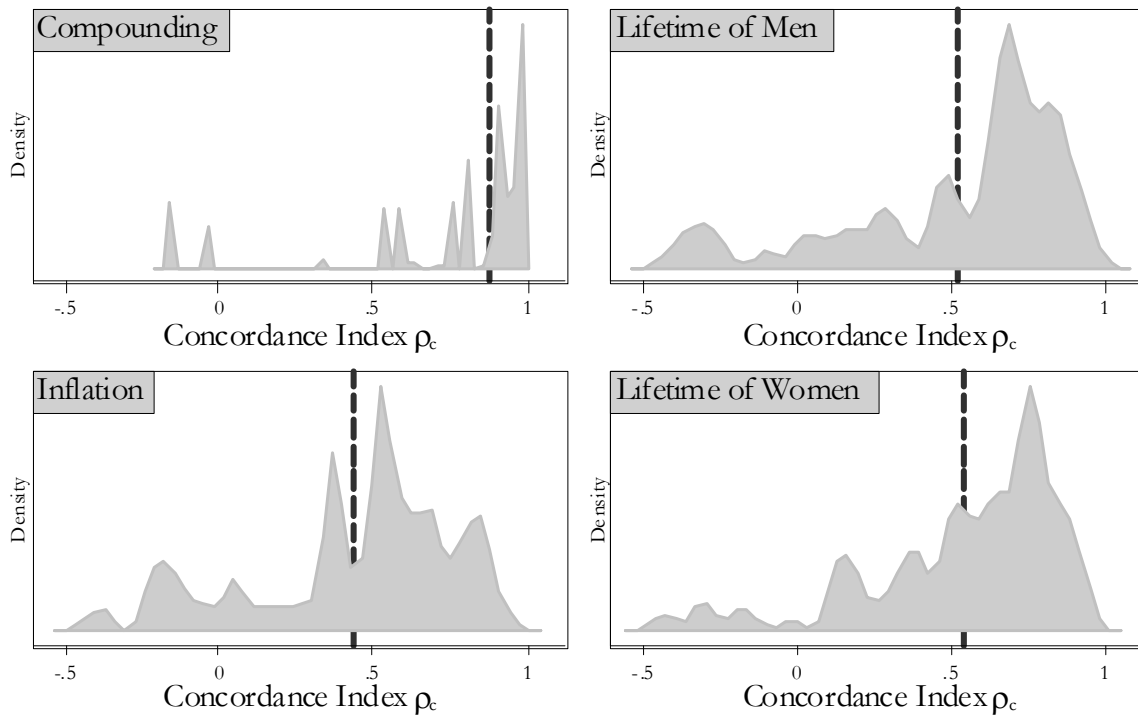


Table 1: Ordered Logit Model of Literacy Index

Variable	Estimate	Standard Error	p-value	95% Confidence Interval
<i>A. Effect on Probability of Being Illiterate</i>				
Inflation	0.10	0.014	<0.001	0.07 ↔ 0.13
Lifetime of men	0.13	0.013	<0.001	0.11 ↔ 0.16
Lifetime of women	0.16	0.013	<0.001	0.14 ↔ 0.19
Female	0.07	0.011	<0.001	0.04 ↔ 0.09
Age, standardized	-0.05	0.007	<0.001	-0.06 ↔ -0.03
Single	-0.15	0.020	<0.001	-0.18 ↔ -0.11
White	0.04	0.016	0.028	-0.004 ↔ 0.07
Finance major	-0.05	0.010	<0.001	-0.07 ↔ -0.03
Non-EU citizen	0.10	0.017	<0.001	0.07 ↔ 0.13
Current smoker	0.02	0.014	0.165	-0.01 ↔ 0.05
Cognitive reflection test	-0.004	0.005	0.479	-0.01 ↔ 0.01
Berlin numeracy test	-0.10	0.005	<0.001	-0.11 ↔ -0.09
<i>B. Effect on Probability of Being Semi-Literate</i>				
Inflation	-0.02	0.004	<0.001	-0.03 ↔ -0.01
Lifetime of men	-0.03	0.004	<0.001	-0.04 ↔ -0.02
Lifetime of women	-0.04	0.004	<0.001	-0.05 ↔ -0.03
Female	-0.01	0.001	<0.001	-0.01 ↔ -0.01
Age, standardized	0.01	0.001	<0.001	0.005 ↔ 0.01
Single	0.04	0.006	<0.001	0.03 ↔ 0.05
White	-0.01	0.003	0.037	-0.01 ↔ 0
Finance major	0.01	0.001	<0.001	0.005 ↔ 0.01
Non-EU citizen	-0.01	0.002	<0.001	-0.02 ↔ -0.008
Current smoker	-0.003	0.002	0.193	-0.01 ↔ 0.002
Cognitive reflection test	0.001	0.001	0.478	-0.001 ↔ 0.002
Berlin numeracy test	0.02	0.001	<0.001	0.01 ↔ 0.02
<i>C. Effect on Probability of Being Literate</i>				
Inflation	-0.08	0.011	<0.001	-0.10 ↔ -0.06
Lifetime of men	-0.10	0.010	<0.001	-0.12 ↔ -0.08
Lifetime of women	-0.13	0.009	<0.001	-0.14 ↔ -0.11
Female	-0.06	0.010	<0.001	-0.08 ↔ -0.04
Age, standardized	0.04	0.006	<0.001	0.03 ↔ 0.05
Single	0.11	0.014	<0.001	0.08 ↔ 0.14
White	-0.03	0.013	0.027	-0.04 ↔ 0.003
Finance major	0.05	0.008	<0.001	0.03 ↔ 0.06
Non-EU citizen	-0.09	0.016	<0.001	-0.12 ↔ -0.06
Current smoker	-0.02	0.012	0.160	-0.04 ↔ 0.01
Cognitive reflection test	0.003	0.004	0.479	-0.01 ↔ 0.01
Berlin numeracy test	0.08	0.005	<0.001	0.07 ↔ 0.09

Figure 9: Concordance Indices

Dashed line is the average



References

- Allen, James, *Natural Language Understanding* (Redwood City, CA: Benjamin/Cummings Publishing, Second Edition, 1995).
- Andersen, Steffen; Fountain, John; Harrison, Glenn W., and Rutström, E. Elisabet, “Estimating Subjective Probabilities,” *Journal of Risk & Uncertainty*, forthcoming 2014.
- Basu, Laushik, and Foster, James E., “On Measuring Literacy,” *Economic Journal*, 108, November 1998, 1733-1749.
- Bateman, Hazel; Eckert, Christine; Geweke, John; Louviere, Jordan; Thorp, Susan, and Satchell, Stephen E., “Financial Competence and Expectations Formation: Evidence from Australia,” *Economic Record*, 88, March 2012, 39-63.
- Clark, Herbert H., *Arenas of Language Use* (Chicago: University of Chicago Press, 1992).
- Cokely, Edward T.; Galesic, Mirta; Schulz, Eric; Ghazal, Saima and Garcia-Retamero, Rocio, “Measuring Risk Literacy: The Berlin Numeracy Test,” *Judgment and Decision Making*, 7(1), January 2012, 25-47.
- Elder, Todd E., “The Predictive Validity of Subjective Mortality Expectations: Evidence from the Health and Retirement Study,” *Demography*, 50, 2013, 569-589.
- Frederick, Shane, “Cognitive Reflection and Decision Making,” *Journal of Economic Perspectives*, 19(4), 2005, 25-42.
- Grice, Paul, *Studies in the Way of Words* (Cambridge, MA: Harvard University Press, 1989).
- Harrison, Glenn W, Martínez-Correa, Jimmy; Swarthout, J. Todd, and Ulm, Eric “Scoring Rules for Subjective Probability Distributions,” *Working Paper 2012-10*, Center for the Economic Analysis of Risk, Robinson College of Business, Georgia State University, 2012.
- Holt, Charles A., and Smith, Angela M., “An Update on Bayesian Updating,” *Journal of Economic Behavior & Organization*, 69, 2009, 125-134.
- Kadane, J. B. and Winkler, Robert L., “Separating Probability Elicitation from Utilities,” *Journal of the American Statistical Association*, 83(402), 1988, 357-363.
- Karni, Edi, “A Mechanism for Eliciting Probabilities,” *Econometrica*, 77(2), March 2009, 603-606.
- Köszegi, Botond, and Rabin, Matthew, “Revealed Mistakes and Revealed Preferences,” in A. Caplin and A. Schotter (eds.), *The Foundations of Positive and Normative Economics: A Handbook* (New York: Oxford University Press, 2008).
- Kruger, Justin, and Dunning, David, “Unskilled and Unaware of It: How Difficulties in Recognizing One’s Own Incompetence Lead to Inflated Self-Assessments,” *Journal of Personality and Social Psychology*, 77(6) 1999, 1121-1134.
- Lin, Lawrence I-Kuei, “A concordance correlation coefficient to evaluate reproducibility,” *Biometrics*, 45, 1989, 255–268.

- Lin, Lawrence I-Kuei, "A note on the concordance correlation coefficient," *Biometrics*, 56, 2000, 324–325.
- Lusardi, Annamaria, and Mitchell, Olivia S., "Baby Boomer Retirement Security: The Roles of Planning, Financial Literacy, and Housing Wealth," *Journal of Monetary Economics*, 54, 2007, 205-224.
- Lusardi, Annamaria, and Mitchell, Olivia S., "Planning and Financial Literacy: How do Women Fare?" *American Economic Review (Papers & Proceedings)*, 98(2), 2008, 413-417.
- Lusardi, Annamaria, and Mitchell, Olivia S., *Financial Literacy: Implications for Retirement Security and the Financial Marketplace* (New York: Oxford University Press, 2012).
- Lusardi, Annamaria, and Mitchell, Olivia S., "The Economic Importance of Financial Literacy: Theory and Evidence," *Journal of Economic Literature*, 52(1), 2014, 5-44.
- Lusardi, Annamaria; Mitchell, Olivia S., and Curto, Vilsa, "Financial Literacy Among the Young," *Journal of Consumer Affairs*, 44(2), Summer 2010, 358-380.
- Matheson, James E., and Winkler, Robert L., "Scoring Rules for Continuous Probability Distributions," *Management Science*, 22(10), June 1976, 1087-1096.
- Merkle, Christoph, and Weber, Martin, "True Overconfidence: The Inability of Rational Information Processing to Account for Apparent Overconfidence," *Organizational Behavior and Human Decision Processes*, 116, 2011, 262-271.
- Moore, Don A., and Healy, Paul J., "The Trouble With Overconfidence," *Psychological Review*, 115(2), 2008, 502-517.
- Nickerson, Carol A. E., "A Note on 'A Concordance Correlation Coefficient to Evaluate Reproducibility'," *Biometrics*, 53, December 1997, 1503-1507.
- Perozek, Maria, "Using Subjective Expectations to Forecast Longevity: Do Survey Respondents Know Something We Don't Know?" *Demography*, 45(1), February 2008, 95-113.
- Remund, David L., "Financial Literacy Explicated: The Case for a Clearer Definition in an Increasingly Complex Economy," *Journal of Consumer Affairs*, 44(2), Summer 2010, 276-295.
- Savage, Leonard J., "Elicitation of Personal Probabilities and Expectations," *Journal of American Statistical Association*, 66, December 1971, 783-801.
- Savage, Leonard J., *The Foundations of Statistics* (New York: Dover Publications, 1972; Second Edition).
- Shefrin, Hersh, *A Behavioral Approach to Asset Pricing* (New York: Academic Press, Second Edition, 2008).
- Smith, V. Kerry; Taylor, Donald H., and Sloan, Frank A., "Longevity and Expectations: Can People Predict Their Own Demise?" *American Economic Review*, 91, September 2001, 1126-1134.
- van Rooij, Maarten; Lusardi, Annamaria and Alessie, Rob, "Financial Literacy and Stock Market Participation," *Journal of Financial Economics*, 101(2), August 2011, 449–472.

Appendix A: Experimental Instructions (NOT FOR PUBLICATION)

Text in square brackets is for the experimenter only.

A. General Instructions

Introduction

You are now participating in a decision-making experiment. Based on your decisions in this experiment, you can earn money that will be paid to you in cash today. It is important that you understand all instructions before making your choices in this experiment.

Please turn to silent, and put away, your mobile phone, laptop computer, or any other device you may have brought with you. Please do not talk with others seated nearby for the duration of the experiment. If at any point you have a question, please raise your hand and we will answer you as soon as possible.

The experiment consists of a demographic survey and two decision-making tasks. You have already earned £5 for agreeing to participate in the experiment, which will be paid in cash at the end of the session. In addition to this show-up fee, you may earn considerably more from your choices in the decision-making tasks. These tasks and the potential earnings from them will be explained in detail as we proceed through the session.

Before we begin the experiment, we will give you an informed consent form. This form explains your rights as a participant in the experiment, what the experiment is about and how we make payments to you. [Give the informed consent form to subjects and read it out loud.]

We will begin the experiment by asking you to answer some demographic questions. Your answers to those questions will not influence your payoffs. [Give the questionnaire to subjects.]

We will now continue with the first decision-making task. You will be given written instructions, but make all decisions on the computer in front of you. We will distribute the instructions and then read it out loud. Please remain silent during the experiment, and simply raise your hand if you have any question so that an experimenter will come to you.

[Give the first set of instructions to subjects and read it out loud.]
[Determine earnings for each subject.]

We will now continue with the second decision-making task. You will again be given written instructions and make all decisions on the computer in front of you. We will distribute the instructions and then read it out loud. Please remain silent during the experiment, and simply raise your hand if you have any question so that an experimenter will come to you.

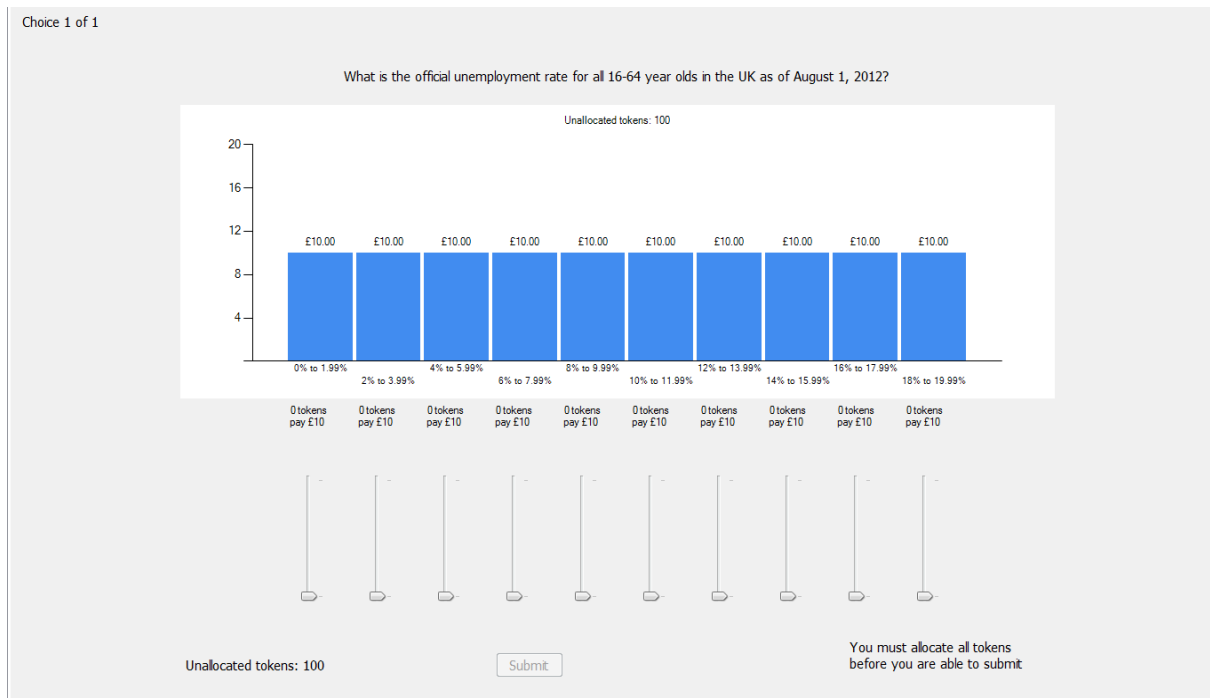
[Give the second set of instructions to subjects and read it out loud.]
[Determine earnings for each subject.]

B. Belief Elicitation Instructions

Your Beliefs

This is a task where you will be paid according to how accurate your beliefs are about certain things. You will be presented with 8 questions of the type we will explain below. You will actually get the chance to play one question presented to you, so you should think carefully about your answer to each question.

Here is an example of what the computer display of such a question might look like.

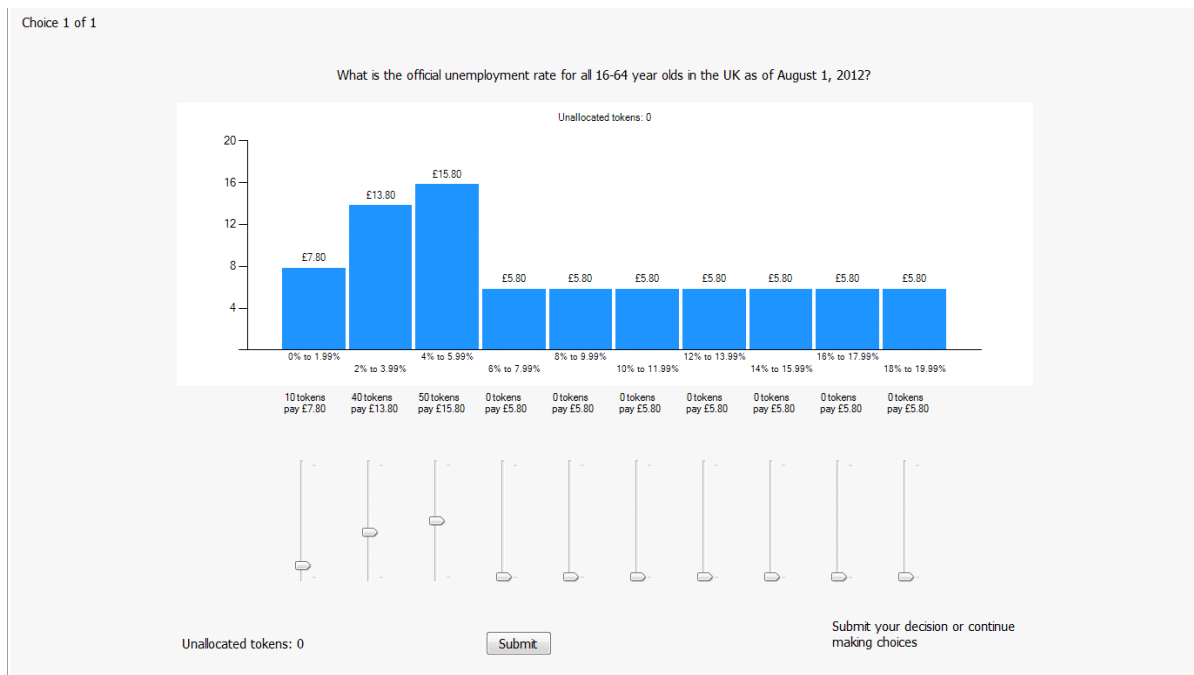


The display on your computer will be larger and easier to read. You have 10 sliders to adjust, shown at the bottom of the screen, and you have 100 tokens to allocate. Each slider allows you to allocate tokens to reflect your belief about the answer to this question. You must allocate all 100 tokens in order to submit your decision, and in this example we start with 10 tokens allocated to each slider. The payoffs shown on the screen only apply when you allocate all 100 tokens. As you allocate tokens, by adjusting sliders, the payoffs displayed on the screen will change. Your earnings are based on the payoffs that are displayed after you have allocated all 100 tokens.

You can earn up to £20 in this task.

Where you position each slider depends on your beliefs about the correct answer to the question. In the above example the tokens you allocate to each bar will naturally reflect your beliefs about the official unemployment rate for all 16-64 years old in the UK as of August 1, 2012. The first bar here corresponds to your belief that the unemployment rate is between 0% and 1.99%. The second bar corresponds to your belief that the unemployment rate is between 2% and 3.99%, and so on. Each bar here shows the amount of money you earn if the official unemployment rate is in the interval shown under the bar.

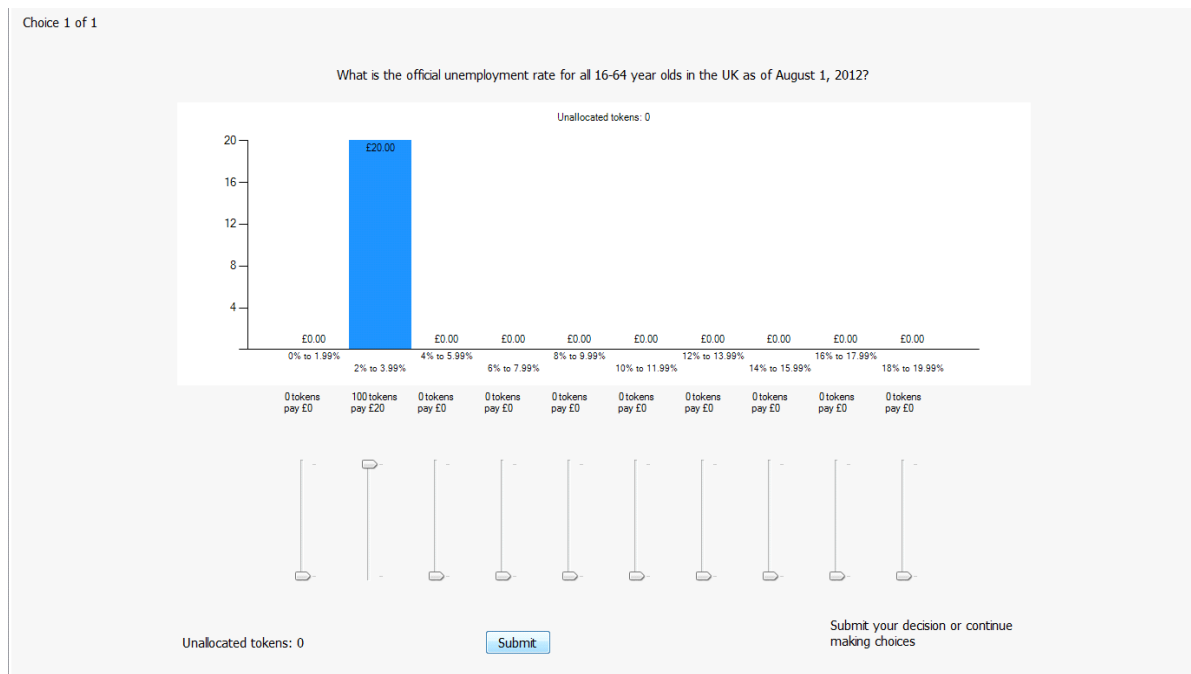
To illustrate how you use these sliders, suppose you think there is a fair chance the unemployment rate is just under 5%. Then you might allocate the 100 tokens in the following way: 50 tokens to the interval 4% to 5.99%, 40 tokens to the interval 2% to 3.99%, and 10 tokens to the interval 0% to 1.99%. So you can see in the picture below that if indeed the unemployment rate is between 4% to 5.99% you would earn £15.80. You would then earn less than £15.80 for any other outcome. You would earn £13.80 if the unemployment rate is between 2% and 3.99%, £7.80 if it is between 0% and 1.99%, and for any other unemployment rate you would earn £5.80.



You can adjust the allocation as much as you want to best reflect your personal beliefs about the unemployment rate.

Your earnings depend on your reported beliefs and, of course, the true answer. For instance, suppose you allocated your tokens as in the figure shown above. The true unemployment is 7.8%. So if you had reported the beliefs shown above, you would have earned £5.80.

Suppose you had put all of your eggs in one basket, and for example allocated 100 tokens to the interval corresponding to 2.5%. Then you would have faced the earnings outcomes shown below.



Note the “good news” and “bad news” here. If the unemployment rate is indeed between 2% and 3.99%, you earn the maximum payoff, shown here as £20. But the true unemployment rate is 7.8%, so you would have earned nothing in this task.

It is up to you to balance the strength of your personal beliefs with the risk of them being wrong. There are three important points for you to keep in mind when making your decisions:

- **Your belief about the correct answer to each question is a personal judgment that depends on the information you have about the different events.**
- **Depending on your choices and the correct answer you can earn up to £20.**
- **Your choices might also depend on your willingness to take risks or to gamble.**

The decisions you make are a matter of personal choice. Please work silently, and make your choices by thinking carefully about the questions you are presented with.

When you are happy with your decisions, you should click on the **Submit** button and confirm your choices. When everyone is finished we will reveal the right answer to each of the 8 questions. Then an experimenter will come to you and ask you to roll a 10-sided die until a number between 1 and 8 comes up to determine which question will be played out. The experimenter will record your earnings according to the correct answer and the choices you made.

All payoffs are in cash, and are in addition to the £5 show-up fee that you receive just for being here.

Are there any questions?

C. Demographic and Other Hypothetical Questions

ID _____

In this survey most of the questions asked are descriptive. We will not be grading your answers and your responses are completely confidential. Please think carefully about each question and give your best answers.

1. What is your age? _____ years

2. What is your sex? (Circle one number.)

01 Male

02 Female

3. Which of the following categories best describes you? (Circle one number.)

01 White

02 Mixed

03 Asian or Asian British

04 Chinese or other ethnic group

05 Prefer not to say

4. What is your main field of study? (Circle one number.)

01 Accounting

02 Economics

03 Finance

04 Business Administration, other than Accounting, Economics, or Finance

05 Education

06 Engineering

07 Health and Medicine

08 Biological and Biomedical Sciences

09 Math, Computer Sciences, or Physical Sciences

10 Social Sciences or History

11 Law

12 Psychology

13 Modern Languages and Cultures

14 Other Fields

5. What is your year of studies? (Circle one number.)

01 First year

02 Second year

03 Third year

04 Masters

05 Doctoral

6. What is the highest level of education you expect to complete? (Circle one number)

- 01 Bachelor's degree
- 02 Master's degree
- 03 Doctoral degree
- 04 Professional qualification

7. As a percentage, what is your current average mark if you are doing a Bachelor's degree, or what was it when you did a Bachelor's degree? This mark should refer to all your years of study for this degree, not just the current year. Please pick one by rounding up or down to the nearest number:

- 01 Above 70%
- 02 Between 60 – 69%
- 03 Between 50 – 59%
- 04 Between 40 – 49%
- 05 Less than 40%
- 06 Have not taken courses for which grades are given.

8. What is your citizenship status?

- 01 British Citizen
- 02 EU Citizen (non-British Citizen)
- 03 Non-EU Citizen

9. Are you currently:

- 01 Single and never married?
- 02 Married?
- 03 Separated, divorced or widowed?

10. How many people live in your household? Include yourself, your spouse and any dependents. Do not include your parents or roommates unless you claim them as dependents.

11. Please circle the category below that describes the total amount of income before tax earned in the calendar year 2007 by the people in your household (as "household" is defined in question 10). [Consider all forms of income, including salaries, tips, interest and dividend payments, scholarship support, student loans, parental support, social security, alimony, and child support, and others.]

- 01 Less than £10,000
- 02 £10,000 – £19,999
- 03 £20,000 – £29,999
- 04 £30,000 – £49,999
- 05 Over £50,000

12. Please circle the category below that describes the total amount of income before tax earned in the calendar year 2007 by your parents. [Consider all forms of income, including salaries, tips, interest and dividend payments, social security, alimony, and child support, and others.]

- 01 Less than £10,000
- 02 £10,000 – £19,999
- 03 £20,000 – £29,999
- 04 £30,000 – £49,999
- 05 Over £50,000
- 06 Don't Know

13. Do you currently smoke cigarettes? (Circle one number.)

- 00 No
- 01 Yes

If yes, approximately how much do you smoke in one day? _____ cigarettes.

14. A bat and a ball cost £1.10 in total. The bat costs £1.00 more than the ball. How much does the ball cost? £_____

15. If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets? _____ minutes

16. In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake? _____ days

17. Imagine we are throwing a five-sided die 50 times. On average, out of these 50 throws how many times would this five-sided die show an odd number (1, 3 or 5)? _____ out of 50 throws.

18. Out of 1,000 people in a small town 500 are members of a choir. Out of these 500 members in the choir 100 are men. Out of the 500 inhabitants that are not in the choir 300 are men. What is the probability that a randomly drawn man is a member of the choir? (Please indicate the probability in percent).
_____%

19. Imagine we are throwing a loaded die (6 sides). The probability that the die shows a 6 is twice as high as the probability of each of the other numbers. On average, out of these 70 throws, how many times would the die show the number 6? _____ out of 70 throws.

20. In a forest 20% of mushrooms are red, 50% brown and 30% white. A red mushroom is poisonous with a probability of 20%. A mushroom that is not red is poisonous with a probability of 5%. What is the probability that a poisonous mushroom in the forest is red? _____%

Appendix B: Detailed Statistical Results (NOT FOR PUBLICATION)

Table B1: Estimates for Interest Compounding Question

(True response was £110.4)

	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
female	.1208637	.1390905	0.87	0.385	-.1517486	.393476
ageS	.3968186	.1936625	2.05	0.040	.017247	.7763901
single	.6368908	.4479348	1.42	0.155	-.2410453	1.514827
white	.2266173	.1577412	1.44	0.151	-.0825497	.5357843
finance	.1057089	.1583877	0.67	0.505	-.2047253	.416143
non_eu	-.0022251	.127354	-0.02	0.986	-.2518344	.2473843
smoker	-.315451	.3116881	-1.01	0.312	-.9263484	.2954465
crtS	-.0427266	.0771227	-0.55	0.580	-.1938843	.1084311
berlinS	.0508408	.0829787	0.61	0.540	-.1117944	.213476
_cons	109.71	.4570172	240.06	0.000	108.8142	110.6057
sigma	.9485433	.1233048			.7352012	1.223793

Table B2: Estimates for Inflation Question

(True response was £198)

	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
female	.6555166	.30101	2.18	0.029	.0655479	1.245485
ageS	-.049693	.2060632	-0.24	0.809	-.4535694	.3541834
single	-.0346816	.4901256	-0.07	0.944	-.9953101	.9259469
white	-.1477976	.4230715	-0.35	0.727	-.9770024	.6814072
finance	-.3338726	.2534093	-1.32	0.188	-.8305456	.1628005
non_eu	-.7831535	.4698084	-1.67	0.096	-1.703961	.1376539
smoker	-.095918	.3464372	-0.28	0.782	-.7749224	.5830864
crtS	.0795068	.1481538	0.54	0.592	-.2108692	.3698829
berlinS	.0675041	.1527122	0.44	0.658	-.2318064	.3668145
_cons	199.1655	.601727	330.99	0.000	197.9862	200.3449
sigma	1.696348	.1015747			1.508504	1.907584

Table B3: Estimates for Remaining Lifetime of Men Question

(True response was 59.1 years)

	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
female	2.262256	2.497932	0.91	0.365	-2.633602	7.158114
ageS	2.446757	2.312314	1.06	0.290	-2.085295	6.978809
single	-3.488838	7.019735	-0.50	0.619	-17.24727	10.26959
white	7.081467	3.446061	2.05	0.040	.3273123	13.83562
finance	3.842162	2.688428	1.43	0.153	-1.427061	9.111385
non_eu	1.31051	3.622383	0.36	0.718	-5.789229	8.41025
smoker	.396316	2.321885	0.17	0.864	-4.154495	4.947126
crtS	-.9155401	1.338364	-0.68	0.494	-3.538686	1.707606
berlinS	1.790858	1.455642	1.23	0.219	-1.062149	4.643865
_cons	48.28431	7.496065	6.44	0.000	33.5923	62.97633
sigma	15.8356	.8974917			14.17073	17.69607

Table B4: Estimates for Remaining Lifetime of Women Question

(True response was 62.9 years)

	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
female	.1826926	2.753928	0.07	0.947	-5.214907	5.580292
ageS	4.212811	1.999146	2.11	0.035	.2945579	8.131064
single	4.533479	5.629901	0.81	0.421	-6.500924	15.56788
white	4.043839	3.208968	1.26	0.208	-2.245623	10.33333
finance	2.796309	2.418511	1.16	0.248	-1.943885	7.536503
non_eu	-.0591597	3.629117	-0.02	0.987	-7.172098	7.053779
smoker	1.303762	2.747118	0.47	0.635	-4.08049	6.688014
crtS	-.3553245	1.254384	-0.28	0.777	-2.813871	2.103222
berlinS	.4161316	1.42578	0.29	0.770	-2.378345	3.210608
_cons	47.63064	7.041492	6.76	0.000	33.82957	61.43171
sigma	15.29233	.7719308			13.8518	16.88265