WORKING PAPER

Eliciting Subjective Expectations in Internet Surveys

ADELINE DELAVANDE, SUSANN ROHWEDDER

WR-589

June 2008

This paper series made possible by the NIA funded RAND Center for the Study of Aging (P30AG012815) and the NICHD funded RAND Population Research Center (R24HD050906).

RAND LABOR AND POPULATION

Eliciting Subjective Expectations in Internet Surveys

Adeline Delavande[*,♦] and Susann Rohwedder[*]

[*]RAND
[♦]Universidade Nova de Lisboa and CEPR

June 2008

JEL Codes: C42, C81, D83, D84, D91

Acknowledgements
This research was supported by a grant from the National Institute on Aging (P01AG008291). Rohwedder is grateful for additional funding from the National Institute on Aging provided under grant R03AG024269. We would like to thank Norbert Schwarz and other participants of the 2007 conference on "Subjective Probabilities and Expectations: Methodological Issues and Empirical Application to Economic Decision-Making" in Jackson, Wyoming for helpful comments.

Abstract

Individuals' subjective expectations are important in explaining heterogeneity in individual choices, but their elicitation poses some challenges, in particular when one is interested in the subjective probability *distribution* of an individual. We have developed an innovative visual representation for Internet surveys that has several advantages over previously used formats. In this paper we present our findings from testing this visual representation in the context of individuals' Social Security expectations. Respondents are asked to allocate a total of 20 balls across seven bins to express what they believe the chances to be that their future Social Security benefits would fall into any one of those bins. Our data come from the Internet Survey of respondents to the Health and Retirement Study which is representative of the U.S. population age 51 and older. To contrast the results from the visual format with a previously used format we divided the sample into two random groups and administered both, the visual format and the more standard percent chance format. Our findings suggest that the main advantage of the visual format is that it generates usable answers for virtually all respondents in the sample while in the percent chance format a significant fraction (about 20 percent) of responses is lost due to inconsistencies. Across various other dimensions the visual format performs similarly to the percent chance format, leading us to conclude that the bins-and-balls format is a viable alternative that leads to more complete data.

Introduction

Studies of individual decision making in a dynamic setting need to take into account individuals' expectations about future events. Traditionally, to close the model, researchers have met this need by making assumptions about individuals' expectations or the expectation formation process. This shortcut is largely due to the lack of adequate data. Over the last 15 years there have been some important advances in the collection of information on subjective expectations, most importantly the move from a deterministic survey question format to one that captures some of the uncertainty necessarily associated with an event that has not yet materialized.[1] Several general-purpose surveys now collect expectations data across several domains in the form of subjective probabilities rather than point expectations.[2] Questions usually take the "percent chance" format

> *"On a scale from 0 to 100 where 0 means you think there is no chance and 100 means that you are absolutely certain, what do you think are the chances that [...]"*

Subjective probabilities collected in this manner have been shown to vary systematically with covariates in the same way as the actual outcomes, to be predictive of actual outcomes and of actual and expected economic behavior (e.g., Dominitz 1998, Hurd and McGarry 1995, 2002, Gan et al. 2004, Delavande and Willis 2008, Delavande forthcoming). However, this question format tends to produce focal answers, in particular bunching of responses at 50 percent, but also at 0 and 100. Bruine de Bruin et al. (2002) find that the fraction of focal answers is reduced when supplying a visual representation such as a response scale that explicitly shows the numeric

---

[1] For an overview of and the state of knowledge of expectations data, see Manski (2004).

[2] Examples of large household surveys that have collected data on individuals' subjective expectations are the Health and Retirement Study, the Panel Study of Income Dynamics, the National Longitudinal Survey of Youth, the Michigan Survey of Consumers, the Survey of Economic Expectations, and the Survey of Health, Ageing and Retirement in Europe. See for example Hurd and McGarry (1995), Dominitz and Manski (1997, 2004, 2007) and Fischhoff, et al. (2000) for analyses of these expectations data.

response options from 0 to 100. Internet-based interviews open the possibility to much richer visual representations, including tailoring to individual respondents' previous answers. We have designed a survey instrument that takes advantage of these features of the Internet and that goes beyond asking just about one point on individuals' subjective probability distribution. Using an innovative visual representation we elicit information about individuals' entire subjective *distribution* of beliefs in the context of individuals' expectations about their future Social Security benefits. We present respondents with seven bins each of which represents a range into which an individual's future Social Security benefits might fall. Respondents are asked to allocate a total of 20 balls across those seven bins to express what they believe the chances to be that their future Social Security benefits would fall into any one of those bins.

We are not the first to elicit information from the entire *distribution* of individuals' expected Social Security benefits. Dominitz and Manski (2006) have done so in the Survey of Economic Expectations using the percent chance format. They ask about six different points on the cumulative distribution function. Even though the format is somewhat repetitive for respondents they achieved a response rate of 97 percent among those who provided a minimum and a maximum Social Security benefit amount.[3] Intuitively, the bins-and-balls format has several advantages: first, it avoids the repetitive series of questions inherent in the percent-chance format (due to asking about several different thresholds); second, it yields internally consistent data (no violations of monotonicity of the cumulative distribution function); third, the information comes in a form that resembles a probability density which may help the respondents and,

[3] In Dominitz and Manski (2006), only respondents who provided a smallest and a highest possible value of expected Social Security benefits were asked the six follow-up questions to elicit several points on individuals' subjective probability distribution. 97 percent of the sample answered these follow-up questions; however, a large fraction of the total sample did not provide the minimum or maximum value resulting in a significant reduction in usable answers.

finally, by increasing the number of bins, it can provide more detail about the probability distribution in just one question.

To test our bins-and-balls format against the percent chance format, we randomized our sample in two groups and administered a different format to each one of them. Our data come from the Internet Survey of respondents to the Health and Retirement Study (HRS) which is a longitudinal survey that is representative of the U.S. population age 51 and older. We link the observations on Social Security expectations from the Internet Survey to the rich background information collected on the same respondents in the main HRS survey. We can therefore investigate the informational content of the Social Security expectations; find how the uncertainty in expected Social Security varies by individual characteristics and whether the bins-and-balls format differs in this respect from the percent-chance format. For example, the dispersion of Social Security expectations should be larger for individuals who report a higher risk of losing their job or who are uncertain about the timing of their retirement. To the extent that this relationship and other similar ones appear to be stronger in one question format than the other would be an indicator of higher data validity.

In this paper we present a thorough assessment of the elicited Social Security expectations that we collected and contrast the visual format with patterns found for the percent chance format. Our objective is to find whether and to what extent the visual format – rendered possible by the Internet mode – is a viable alternative for eliciting the distribution of beliefs and to what extent it might even be advantageous compared to the percent chance format. Our findings suggest that the main benefit of the visual format is that it generates usable answers for virtually all respondents in the sample while in the percent chance format a significant fraction (20 percent) of responses is lost due to inconsistencies. Along other dimensions the two formats yield patterns that are closely comparable. Item non-response is very low for both formats and statistics of central tendency of the distribution of Social Security benefits are very similar. However, the shape of the distributions is different. The visual format tends to attract more probability mass in the middle than the percent chance format. Based on additional experiments we find that this does not appear to be due to anchoring bias in the visual format. When relating measures of uncertainty about future Social Security benefits to covariates such as uncertainty in other domains, both formats show systematic variation of the anticipated sign.

1 Data

1.1. The HRS Internet Survey

The data for our study come from a module of the HRS Internet Survey which is a supplementary survey of the Health and Retirement Study (HRS). The HRS is a panel survey that is representative of the U.S. population age 51 and over. A sub-sample of about 6,000 respondents is eligible to participate in the HRS Internet Survey. We use data from the second wave which consists of HRS respondents who reported having Internet access in the HRS core survey in 2006. It was collected in two phases with the first part of the sample (38 percent) being interviewed in the early part of 2006 (Phase 1) and the second part of the sample (62 percent) being interviewed in the summer of 2007.[4] The unit response rate was 70 percent in both phases. Most of our analytical sample comes from phase 2, but we bring in some data from phase 1 (see below for further detail). With Internet use being the main selection criterion the resulting sample is more educated than the general population.

[4] The reason for postponing the interview to a later second phase for one group of the sample was that this group had been previously assigned (at random) to participate in another supplemental study and would have had three HRS-related interviews within months of each other had their internet interview not been postponed.

1.2 Experimental design

The survey questions that we focus on in this study are embedded in a sequence eliciting several aspects related to future Social Security receipt: whether the person expects to receive any Social Security benefits in the future, at what age and how much. To elicit the subjective distributions of beliefs about future Social Security benefits, respondents are randomized into one of two different formats: (*i*) a new visual format that we designed and (*ii*) the standard percent chance format.[5]

1.2.1 Question wording and experimental design of lead-up questions

Respondents who reported not currently receiving Social Security benefits are asked the percent chance that they will receive Social Security benefits some time in the future. Those who provide a positive probability are asked about their subjective expectations regarding their claiming age and future Social Security benefits (conditional on receiving any). Before providing information on the subjective distributions of beliefs about their Social Security benefits, respondents are first asked to give a point estimate of their expected Social Security benefit amount:

> *How much do you expect your monthly Social Security benefits to be in today's dollars?*

We use this point estimate to tailor the thresholds for questions about individuals' subjective distributions about future benefits which follow next.

[5] Respondents are also randomized into whether they are asked about their future Social Security benefits conditional or unconditional on their self-reported expected claiming age. In earlier work (Delavande and Rohwedder, 2007) we show that conditioning on expected claiming age leads to distributions that are not statistically different from each other. This is because the respondents' underlying distribution of expected claiming ages is almost universally symmetric around the point estimate of the expected claiming age. As a result, we pool respondents who were asked their expectations conditional or unconditional on claiming age in this paper.

1.2.2 Subjective distributions of Social Security benefit amounts: Visual format

The visual format elicits information on individuals' expectations about their future Social Security benefits in a way that mimics the density function of their subjective beliefs. The introduction includes a couple of examples to familiarize respondents with the exercise:

Often people are uncertain about their future Social Security benefits.

In the next question, we ask you to think about what your monthly benefits might be. We will show you 20 balls that you can put in seven different bins, reflecting what you think are the chances out of 20 that your future Social Security benefits fall in each bin. The more likely you think it is that your benefits fall in a given bin, the more balls you should assign to this bin. For example, if you put all the balls in the bin $500 - $800, it means you are certain that the amount you will receive is between $500 and $800. Another example is illustrated on the next screen.

In this example, there are no balls in the ranges below $520 or above $1040, so it is certain that benefits will not be lower than $520 or higher than $1040. 12 out of 20 balls are in the bin $520 - $780 meaning that the chance that benefits are between $520 and $780 is 12 out of 20 (60 percent). There is a smaller chance, 8 out of 20 (40 percent), that benefits are between $780 and $1040.

Note the examples of interpretation that are included to provide guidance also to those who may not be all that comfortable in thinking in terms of percentages.

Figure 1 shows the screen that respondents see when allocating the 20 balls into 7 bins.

1.2.3 Subjective distributions of Social Security benefit amounts: Percent chance format

The percent chance format follows Dominitz and Manski (2006) in requiring respondents to provide multiple points on their cumulative distribution function. In particular, respondents are asked the following four questions:

Often people are uncertain about their future Social Security benefits. In the next few questions we will ask you about the chances that your future Social Security benefits turn out to be higher or lower than certain values.

*Now what about the chances that your Social Security benefits might be **higher**: On a scale from 0 to 100, where 0 means no chance and 100 means you are absolutely certain, what do you think is the percent chance that your Social Security benefits will be more than **$[X3] per month**?*

If answer to previous question>0, then:

*Still about the chances that your Social Security benefits might be **higher**: On the same scale, what do you think is the percent chance that your Social Security benefits will be more than **$[X4] per month**?*

*Now what about the chances that your Social Security benefits might be **lower?**: On a scale from 0 to 100, where 0 means no chance and 100 means you are* *absolutely certain, what do you think is the percent chance that your Social Security benefits will be less than* ***$[X2] per month****?*

If answer to previous question>0, then:

Still about the chances that your Social Security benefits might be ***lower****: On the same scale, what do you think is the percent chance that your Social Security benefits will be less than* ***$[X1] per month****?*

Where $X1 < X2 < X3 < X4$.

Within the percent chance format there is another randomization: half the sample is first asked about the chances that benefits could be higher and then about the chances that benefits might be lower. For the other half the ordering is reversed. The rationale for this randomization is that the "higher-" or "lower-" wording might lead to anchoring bias in the answers which would average out in summary statistics at the population level.

1.2.4 Design of thresholds

For both the visual and percent chance format, the thresholds depend on respondents' reported expected Social Security amount.[6] The motivation for using tailored thresholds is presented in Dominitz and Manski (1997). First, it allows us to ask about thresholds that span a support that is relevant for the respondent, rather than asking about a range that might be too wide or too narrow and thus less informative. Second, it decreases potential anchoring effects where respondents' answers might be influenced by the amounts associated with the thresholds.

We designed the thresholds to be centered around the respondents' point estimate of their expected future Social Security benefits. They are computed in the same way for the visual and

[6] The median expected amount is $1,100 per month.

the percent chance format, though the former has two additional thresholds in the extreme right and left (see Figure 2).[7] When the point estimate is missing, we provided standard thresholds.[8]

1.3 Analytical sample

We focus in this study on respondents' expectations about their future Social Security benefits. Questions about these expectations are only relevant for those respondents who do not receive Social Security benefits at the time of the interview. This is the case for 67 percent of the sample. The earliest age at which an eligible person can claim Social Security benefits is at age 62 and the latest age is 70. As a result the oldest person in our analytical sample is 72. Among those not currently receiving Social Security benefits, 5.0 percent say "zero percent" for the chance to receive Social Security in the future. The remainder of the sample is asked about their Social Security expectations.

Our analytical sample has 1,418 observations.[9] Of these 55 percent were randomized into the bins-and-balls format and 45 into the percent chance format.

While the original sampling frame is the population representative pool of HRS respondents, eligibility for the Internet survey is based on respondents using the Internet on a regular basis. As a result the sample has higher education than the general population. Table 1 shows the distribution of sex, age and education for our analytical sample. For comparison we show the characteristics of HRS 2006 respondents less than age 72 who do not currently receive Social Security, weighted and unweighted. The distributions of sex and age are closely comparable

[7] Dominitz and Manski (2006) also use preliminary questions to determine individual-specific thresholds. However, rather than using the point estimate they ask respondents about the lowest and highest future Social Security benefits amount. We decided to use the point estimate because non-response to these lowest and highest amounts was high in their survey (30 percent).

[8] The standard thresholds are $0, $300, $600, $900, $1,200, $1,500 and $1,800 for the bins-and-balls format, and $600, $900, $1,200, $1,500 for the percent chance format.

[9] This includes 397 respondents from the Phase 1 of data collection. The survey design was the same in Phase 1 and Phase 2 of the second wave of the HRS Internet Survey for the group who answered the visual format. This results in a larger sample size for this group.

between the unweighted whole HRS sample and our analytical sample. However, as expected, our analytical sample has higher education, with 41.5 percent having completed college or above.[10] We do not use weights in our analysis as we are presenting results for randomized groups to make advances in survey methodology rather than population estimates.

2. Basic data quality: Visual versus Percent chance format

2.1. Response Rates

Item non-response rates are very low throughout: about three percent for the point expectations and less than two percent for the subjective distribution of beliefs for both formats. Differences across the two question designs are very small as a result. These minimal differences do not lead to any one of the question designs to stand out over the other. Note that in the percent chance format 98.4 percent answered all four questions about the different thresholds.[11]

2.2. Survey Time

Using observations on the timing of keystrokes we find only a small difference in how long it took respondents to answer the questions in the different formats (84.7 seconds to go through the two introductory screens of the visual format and to allocate the balls and 90.2 seconds to complete the sequence of the four percent chance questions[12]). Note that allocating the balls accounted for only 21 seconds on average . This suggests that if a researcher is interested in eliciting more than one distribution within the same survey, there might be gains in terms of

---

[10] The high fraction of female respondents in our analytical sample may seem striking at first. However, this is driven by the high fraction of female respondents in the HRS core sample and not by the eligibility criteria for the HRS Internet study. In fact males and females are equally likely to use the internet on a regular basis and the willingness to participate in an internet study is virtually the same for both sexes as well.

[11] As a comparison, about 96.8 percent of respondents to the Survey of Economic Expectations answered the *six* thresholds about future Social Security benefits (Dominitz and Manski, 2006).

[12] The medians are 61 and 65 seconds respectively.

survey time from using the bins-and-balls format, as the introduction screens can easily be skipped or shortened when eliciting additional distributions.

2.3. Usable Answers

A common problem with administering a series of percent chance questions to elicit multiple points on respondents' subjective probability distribution about an event is that respondents may not respect the monotonicity across thresholds (i.e. the subjective probability of, say, Social Security benefits being more than $500 is larger than the probability of Social Security benefits exceeding any other higher threshold). Respondents whose answers violate this monotonicity property may not master the concept of probabilities and their answers are difficult to interpret and use in empirical analysis. This is what we call "unusable answers" in the context of the percent chance format. This problem does not arise in the bins-and-balls format. However, in the bins-and-balls setting respondents may fail to allocate all 20 balls. This is what we call "unusable answers" in this format.[13]

More than 97 percent of the answers are usable for the bins-and-balls format. This is in stark contrast to the percent chance format where observations are lost first because some respondents do not answer all the four questions about the four different thresholds (1.6 percent), and second because a considerable fraction of respondents' answers violate the monotonicity property (about 20 percent of those who answered all four questions).[14] This not only reduces the sample size of the percent chance format, but it also introduces selection effects. For example, 68 percent of the

[13] One could also argue that answers where not all balls were allocated across the bins could still be used by rescaling so that whatever the total number of balls would amount to 100 percent probability. However, we use the stricter criterion that all 20 balls have to be allocated.

[14] Dominitz and Manski (2006) ask about the minimum and maximum future Social Security benefits to design the thresholds and inform respondents when their answers violate monotonicity, resulting in 65 percent of their overall sample having usable answers. Note that their results may not be directly comparable because of slight differences in the survey design, but also due to the different survey mode. Their interview is conducted over the phone while ours is self-administered over the Internet.

respondents who violated the monotonicity property attended college, compared to 76 percent among those who provided consistent answers. In addition, those who violated the monotonicity property seem to have more difficulty overall with other probabilistic questions. Using all the expectations questions from HRS 2006, we constructed an index which captures ones' ability to "think probabilistically" in the spirit of Lillard and Willis (2002). The index computes for a particular respondent the proportion of "Don't know", "Refuse" and "50 percent" out of all the expectations questions that are phased using the percent chance format (section P of the HRS core survey instrument). The average index is 20.0 percent among respondents who violated the monotonicity, compared to 17.9 among those who did not.[15]

2.4. Precision of Information

One of the advantages of the bins-and-balls design is that we can present a relatively large set of bins (or intervals) without increasing much the burden to respondents. In contrast, having answers for seven intervals requires asking respondents six percent chance questions. At the same time, however, the bins-and-balls format limits respondents' precision in the allocation of the probability mass in a given bin. In the current design, respondents have to allocate 20 balls, which implies that the smallest positive probability that they can allocate to a bin is 5 percent. The percent chance format allows respondents to report positive probabilities with precision down to the single integer.

We evaluate to what extent respondents use the precision provided by each format. Table 2 presents the total number of bins used by respondents. While respondents randomized into the bins-and-balls format could use up to seven bins, only one percent chose to use more than five

[15] The differences in the fraction who attended college and in the index of probabilistic thinking are statistically significant at the 10-percent level when comparing the two samples.

bins. We also check the proportion of respondents who provide a percent chance which is not a multiple of five. This is the case for less than 3.5 percent of the answers per bin.[16] Overall, very few respondents take advantage of the full precision provided by the format they are given and the limited precision in the bins-and-balls format imposed by the number of balls is not expected to lead to any material difference between the bins-and-balls format and the percent format.

2.5. Shape of the Elicited Distributions

In this section, we compare the shape of the distributions elicited using the two formats to evaluate whether there are systematic differences. To make the data from the two formats directly comparable, we look at the probability mass (on a scale from 0 to 100) distributed across five bins, that is, we aggregate the number of balls in the first and second bin, and in the sixth and seventh bin for the bins-and-balls format. We restrict the analysis to respondents who provided usable answers, according to the definition above.

2.5.1 Most Common Answers

For both formats, the most common answer is to allocate all the probability mass in the middle. However, this answer is more common in the visual format than in the percent chance format (21 percent vs. 10 percent).[17] The second most common answer (six percent of respondents to bins-and-balls) in the visual format is to allocate 50 percent of the probability mass in the middle bin, and 50 percent in the second bin. The third most common answer (five percent of respondents) is to allocate 75 percent of the probability mass to the middle bin, and 25 percent to the fourth bin. For the percent chance format the second and third most common answers are different: two percent of the respondents allocated 50 percent of the probability mass in the

[16] Similar patterns are reported in Manski and Molinari (2008) for the whole HRS.
[17] Respondents with all the probability mass in the middle bin are those who answered 0 percent to all percent chance questions.

middle bin and 50 percent in the fourth bin; while two percent allocated 50 percent of the probability mass in the middle bin and 50 percent in the fifth bin.

One potential drawback of the percent chance format that may be mitigated in a visual format is the tendency of respondents to use focal answers. People who answer “50 percent” may either truly believe that the chance is about half or they may be uncertain (e.g., Bruine de Bruin et al., 2000; Hill et al., 2006). We investigate the pattern of focal answers for the four percent chance questions in Table 3.[18] As expected, zero-percent answers are substantially more frequent for the extreme thresholds (49 and 40 percent) as many respondents are fairly certain that their expected Social Security benefits will not be that far off their previously reported point estimate. Subjective probabilities about the two middle thresholds attract a large proportion of zeros and 50s (about 45 percent of the answers), and there appears to be excessive bunching at 50 percent. The tendency to answer 50 percent for the two middle thresholds may generate less probability mass in the middle bin for the percent chance format.

2.5.2 Central Tendencies and Spread

Table 4 presents the mean and percentiles of the distribution of probability mass by bins for each format. The visual and percent chance format generate similar central tendency. Respondents tend to allocate most probability mass in the middle bin. Yet, the percent chance format generates more spread out distributions: respondents who were administered the percent chance format allocate on average more mass in the two extreme bins and less in the middle bin. On average the middle bin attracts 59.9 percent of the probability mass for the visual format compared to 43.0 in the percent chance format. The first bin attracts on average 2.7 and 9.6 percent of the probability mass respectively. The difference across formats is also reflected in the total number of bins used (see Table 2): respondents in the bins-and-balls format tend to use a

[18] Note that the percent chance questions do not ask directly about the central bin which contains the point estimate.

smaller number of bins than those in the percent chance format. For example, 73 percent of the bins-and-balls format respondents used two bins or less, compared to 32 percent of the percent chance format respondents.

The difference in probability mass in the middle bin across formats is driven in part by two facts: respondents to the visual format are more likely to allocate *all* the probability mass in the middle bin (20% of the respondents vs. 10%) and less likely to allocate *no* probability mass in the middle bin (9% of the respondents vs. 17%). Given that the middle bin contains respondents' own previously reported point estimate, this latter group seems surprisingly high in the percent chance format.

We investigate in more detail whether the two formats generate similar means of the subjective distributions. To construct the individual-specific mean of the elicited subjective distribution, we assume that the probability mass reported by a respondent is uniformly distributed within a bin (see the Appendix 1 for details). Table 5 presents the distribution of these computed means for both formats and for the point expectations and shows that the percentiles of these three distributions are very similar.[19] Despite generating different dispersion and shapes of distributions, the two formats yield distributions with very similar means that are in line with the point expectations.

Note that respondents are more likely to allocate more balls in the middle bin when the width of the bin is large in terms of the range of dollar amount covered: the average probability mass in the middle bin is 54.5 percent among respondents with a width less than \$300, compared to 65.5 percent among respondents with a width above \$300. We do not observe such a striking pattern in

[19] Table 5 excludes 12 respondents who reported a monthly point estimate above \$10,000.

the percent chance format: the average probability masses are 42.2 and 43.7 percent respectively.[20]

2.5.3 Symmetry of the Distributions

Restricting our sample to respondents who did not allocate all the probability mass in one bin Table 6 shows that respondents who were administered the percent chance format are more likely to put the same probability on either side of the central bin, that is, they provide a symmetric distribution (20.5 percent vs. 5.5 percent in the bins-and-balls format).[21]

The higher proportion of asymmetric distributions in the bins-and-balls format may indicate that visualizing the density helps respondents to report the shape they have in mind.

2.6. Anchoring Bias

We have found that the visual format attracts more probability mass in the middle than the percent chance format. We investigate whether this may be due to anchoring toward the middle, possibly because respondents are attracted visually towards the center of the scale and/or possibly because they see all the thresholds at once and again tend towards amounts that are closer to the midpoint of the amounts shown on the screen. First, we look at another question using a similar bins-and-balls format within the same survey; second, we report the results from an experiment that we conducted in another Internet survey specifically to test the anchoring hypothesis for the bins-and-balls format.

[20] The difference for the visual format is statistically significant at the 1-percent level using a t-test, while that of the percent chance is not statistically significant at the 10-percent level. If we regress the probability mass of the middle bin on the width of the interval using a linear regression, the coefficient associated with the width is statistically significant at the 10-percent level for the visual format only.

[21] Note that out of the 20 percent of respondents who provided equal probability on either side, about half gave the same answers to thresholds 1 and 3 and the same answers to thresholds 2 and 4. (e.g. the sequence of answers would be for example 20, 10, 20, 10).

Within the same survey wave of the HRS Internet, respondents are asked to allocate 20 balls into 8 bins to reflect the probability that they would claim Social Security at a certain age. Each bin represents an age or age category (61 or less, 62, 63, 64, 65, 66, 67, 68 or more). The number of bins being even for this question, there is not one single central bin. Instead the two bins for ages 64 and 65 are in the middle. Table 7 presents the distributions of balls by bin and shows that the bins with the largest probability mass are 62 and 65 which are the two most common ages for claiming Social Security. The median number of balls for the bin 64, which is also central, is zero. Therefore, the bins-and-balls format does not systematically generate more probability mass in the middle.

Second, we compare the elicited subjective distributions obtained from two different bins-and-balls formats designed specifically to investigate respondents' tendency to be attracted by the middle bin. These data were collected as part of the American Life Panel (ALP) Internet survey.[22] The main respondents of the ALP are recruited among individuals age 18 and older who are respondents to the Monthly Survey of the University of Michigan's Survey Research Center.[23] In addition to the main respondents, the ALP collects information from a snowball sample. Respondents from the snowball sample are spouses, relatives or co-workers of the main respondents who are referred by them to participate in the survey.

In the 8th wave of the ALP, fielded in 2007, we administered to the whole sample a module on Social Security using the same design as in the HRS Internet survey. In May 2008 (wave 24 of the ALP), we administered a module to respondents from the snowball sample aged 30 to 65 identical to that of the 8th wave, the only difference being the definition of the thresholds in the

[22] The ALP is sponsored by the National Institute of Aging (grant number NIA U01AG009740) and is conducted by the RAND Corporation.
[23] To answer the survey, respondents either use their own computer to log on to the Internet or a provided Web TV. See http://rand.org/labor/roybalfd/american_life.html for details.

bins and balls question. In the 24th wave, the third bin, rather than the middle bin, is centered around the point estimate. Figure 3 compares the two threshold definitions: the second bin in the 24th wave is equivalent to the third bin in the 8th wave, the third bin in the 24th wave is equivalent to the middle bin in 8th wave and so on.

Respondents from the snowball sample are younger than the main ALP respondents.[24] As a result, those two sub-samples tend to have expected Social Security amounts and derived thresholds that differ in their levels. To compare the two designs fielded in the 8th wave and the 24th wave, we thus conduct two separate exercises. We first compare the subjective distributions from the snowball samples of waves 8 and 24. We then compare the distributions for all respondents aged 30 to 64 from wave 8 with that of respondents from wave 24, after having weighted the answers of the 24th wave such that the age distribution of the two samples are identical.[25] Table 8 presents the distributions of the expected Social Security amount in the various samples among respondents who provided usable answers for the bins-and-balls questions, and show that the distributions of our comparison groups are similar.[26]

Table 9 presents the distribution of balls into bins for the two snowball samples. The main finding is that, for the 24th wave, the third bin which contains the point estimate attracts a large proportion of the probability mass despite the fact that it is not located in the middle. The median number of balls is 10 and the average is 9.4. This is extremely close to the number of balls in the middle bin (which contains the point estimate) in the 8th wave sample: 9 at the median and 9.5 at mean. The distributions of balls into the bins immediately to the right and to the left of the bin

[24] For example in the 8th wave, the average age of the snowball respondents was 38.8 compared to 52.0 for the main respondents.

[25] We created weights for respondents of the 24th wave such that the proportion of respondents 30-34, 35-39, 40-44, 45-49, 50-54, 55-59 and 60-64 matches that of the 8th wave.

[26] We cannot reject equality of the mean using a t-test (P-value=0.46) for the two snowball samples. While we could reject at 5% equality of the mean between the overall 8th wave sample and the 24th wave sample, we cannot reject it at 5% once we weight the 24th wave sample.

containing the point estimate are also strikingly similar in the two formats. For example, the middle bin in the 24th wave sample contains on average 3.98 balls, which is very close to the average of 3.44 balls in the 5th bin for the 8th wave sample. If we combine the two formats into 6 comparable bins, we cannot reject equality of the average number of balls for each of the bins at the 5 percent level.

Table 10 presents the same distributions of balls for the weighted snowball sample and wave 8 respondents. The same conclusions hold: among the 24th weighted wave sample, the middle bin does not attract a disproportionate amount of balls, and most of the probability mass is attracted to the bin containing the point estimate (10.2 balls on average in wave 24 vs. 10.8 in wave 8). Again, if we combine the two formats into 6 comparable bins, we cannot reject at 5% equality of the average number of balls for each of the bins.

Overall, this suggests that there is no anchoring bias toward the middle created by the bins-and-balls format. Respondents tend to provide strikingly similar subjective distributions relative to the point estimate when the point estimate is located in the middle bin or in the third bin.

We also investigate anchoring in the percent chance format. Half of those who receive the percent format are asked first about the chances that benefits could be higher and then about the chances that benefits might be lower than the thresholds. For the other half the ordering is reversed (see section 1.2.3). While this does not generate a different pattern of answers for the probability mass lower than a threshold, we find that the average probabilities about the chance that benefits might be higher than certain values differ statistically using a t-test depending on the order. The average answer for the percent chance that benefits might be higher than *X3* is 37 among respondents who are asked this question first, compared to 33 among respondents who are asked this question third (P-value=0.075) and the respective medians are 50 and 30. These averages are 18 and 15 respectively for the percent chance that benefits might be higher than *X4,* the next higher threshold (P-value=0.049), and medians are 10 and 5 respectively. This suggests that earlier questions about the same distribution to the right of the point estimate influence respondents' answers.[27]

2.7. Unfolding bins

The objective behind the bins-and-balls format was to generate data that even in its raw form approximates respondents' subjective probability distributions. However, a large proportion of respondents in the bins-and-balls format used only one or two bins revealing only limited detail about their distributions. Taking advantage of the possibilities offered by the Internet survey mode we administered a follow-up bins-and-balls screen to respondents who provided a concentrated distribution of beliefs, i.e. those who allocated all the balls into one or two adjacent bins. These "unfolding bins" split the initial bin(s) containing all the balls into narrower bins and the respondent is re-asked to place 20 balls into those narrower bins with the following introduction:

> *Thank you for your answer. In order to get more precise information, we have now narrowed the size of the bins. By clicking on the + and - buttons under each bin, please put the 20 balls into the bins such that it reflects what you think are the chances out of 20 that your monthly Social Security benefits fall in each bin.*

The number of new bins introduced varies between 2 and 5 bins depending on the width of the initial bin(s) where all the probability mass was placed. See Appendix 2 for details.

---

[27] Note that we do not find any statistical difference by the ordering of questions in the proportion of respondents respecting the monotonicity property.

Out of the 386 respondents from phase 2 who were randomized into the visual format, 186 were eligible for the unfolding bins.[28] Table 11 presents the distribution of the number of unfolding bins that were administered. Out of those 186 respondents, 174 (that is 93.6 percent) re-allocated the 20 balls into the narrower bins. Table 12 presents the distribution of unfolding bins used by respondents. About 80 percent of those who were presented with two unfolding bins allocated all the balls into one bin. About two thirds of those who were presented with three, four or five unfolding bins used two bins. None of the respondents who were presented with four or five unfolding bins chose to use more than three bins to express their beliefs. Table 13 presents the mean and median of the distribution of balls for each of the bins for the usable answers. Respondents tend to provide again distributions concentrated around their point estimates, with on average more probability mass to the right of the point estimates.

This suggests that respondents are willing to provide more precise answers about their distributions of beliefs if asked. The bins-and-balls format is a very efficient way of doing so because only one additional screen is needed irrespective of the level of additional detail to be elicited. As such this application is particularly well suited to Internet surveys. While a similar follow-up could be done with the percent chance format, this would increase survey time (see discussion in section 2.2) and potentially respondents' burden more importantly.

3. Correlation of distribution of answers with uncertainty in other domains

The plausibility of response patterns and how they vary with other covariates may indicate which format yields higher quality data. We investigate in this section how uncertainty in Social Security benefits correlates with uncertainty about related outcomes. For this analysis we restrict

[28] The unfolding bins were introduced in phase 2. Note that three respondents out of the 386 did not allocate all the 20 balls in the first screen, but one of them placed all the balls in the unfolding bins screen.

our analysis to usable answers and exclude 10 respondents who provided an expected monthly Social Security benefit above $10,000 and who are thus asked about very wide bins.[29]

3.1. Univariate analysis

We focus on two measures of uncertainty in Social Security benefits: the number of bins used by the respondents, and the probability mass allocated in the middle bin. Table 14 presents the average of these two measures according to various characteristics, such as the distance to their expected claiming age or the subjective probability that they would loose their job.

Uncertainty in the timing of future benefit receipt (expected claiming age) is likely to translate into uncertainty about the amount of future Social Security benefits. Also the longer the time to claiming the higher one would think the uncertainty that the individual faces regarding the Social Security benefit amount. For both formats respondents who used more bins to express the distribution of future claiming ages or those who are further away from to their expected claiming age used more bins for the distributions of future Social Security benefits and allocated less probability mass in the central bin. Respondents were asked the probability that they would receive Social Security benefits in the future. Almost half of the respondents reported a probability of 95 percent or more. Table 14 shows that those who report a probability less than 95 percent report a more spread out distribution of future Social Security benefits. A similar effect is found among respondents who provide a probability above 50 percent that a reform would reduce their own benefits. Respondents currently working are asked in the HRS the probability that they would lose their job in the next year. For the visual format, we find that the average number of bins increases monotonically with this probability, and the difference is statistically significant at

[29] Results are very similar when we include those respondents. In total, 12 respondents reported a point estimate above $10,000. However, one violated the monotonicity property and one did not allocate all the 20 balls.

the 1-percent level. The relationship is not monotonic for the percent chance format, with people reporting a probability of loosing one's job between 0 and 10 using more bins on average. Being in poor health may be associated with increased certainty about one's ability to work in the future. In the visual format, we find a monotonic relationship between self-reported health and the number of bins used, with people in excellent health using more bins than people in good or poor health. The relationship is not statistically significant at conventional level and non-monotonic in the percent chance format. Finally, we find that, for the visual format only, respondents who expect to claim Social Security based solely on their own record report a less spread-out distribution of future Social Security benefits than those who expect to claim on their spouse's record or on both their and their spouse's record.

Overall, we find that, for both formats, respondents are more likely to have a more disperse distribution when they face uncertainty in related domains. For some domains, such as health and the probability of loosing one's job, we find the *a priori* expected monotonic effect only for the visual format.

3.2. Multivariate analysis

We next conduct regression analyses using these measures of uncertainty as dependent variables. We also use a constructed standard deviation of the subjective distribution (see Appendix 1 for details) as an additional dependent variable.[30] We employ an ordered probit model when the number of bins used by the respondents is the dependent variable and look at the best linear predictors when the probability mass in the middle and the standard deviation are the dependent variables.

[30] The median standard deviation is 151.5 for the visual format and 284.9 for the percent chance.

Table 15 shows that some of the effects shown in cross-section also hold in a multivariate framework, and hold for the standard deviation as well.[31] For both formats, a larger distance to expected claiming age is still associated with a larger number of bins used by the respondents and a larger standard deviation, and the coefficients are statistically significant at the one-percent level. It is also associated with less probability mass in the central bin, though the effect is statistically significant for the percent chance format only. Similarly, a larger number of bins used for the expected claiming age is associated with more bins for the Social Security benefits. For both formats, the coefficients associated with the subjective probability of reform and of receiving Social Security are statistically significant in most specifications and have the expected sign. The regressions introduce additional independent variables than what was presented in Section 3.1. Working for pay is associated with fewer bins for the percent chance format. Wealth has a statistically significant coefficient for the percent chance format only.[32] Overall, the plausibility of response patterns and how they vary with other covariates suggest that both formats yield high quality data.

The regressions presented in Table 15 are however not without caveat for comparison between the two formats. First, the number of observations is larger in the bins-and-balls format, which is partially due to the initial design and partially due to the loss of 20 percent of the observations in the percent chance format. Caution needs to be taken when comparing the significance levels of coefficients across formats. Second, the sample of usable answers in the

[31] We exclude respondents who did not answer HRS 2006, that is 9 respondents in the bins-and-balls format.

[32] Note that the R-square of the linear regression using the probability mass in the middle bin with the percent chance format is substantially higher than that of the visual format, potentially suggesting that the independent variables explain more of the variation in probability mass in the percent chance format. However, this is mostly driven by the fact that the independent variables include many other expectations questions asked in a percent chance format, and that those tend to be strongly associated with the dependent variable when asked in percent chance. If we remove those, the R-squares of the two formats are basically identical (0.07 for the visual vs. 0.09 in the percent chance format).

percent chance format suffers from selection (see section 2.3), and some of the relationships presented in the regression might be biased as a result.

4. Conclusions

This paper presents the results of an exploratory data collection we have undertaken to elicit subjective expectations using a new visual format made possible in Internet surveys. Our point of departure was that a visual format might elicit more detailed information about individuals' distribution of beliefs without increasing respondent burden, and allow respondents to visualize the density s/he provides, and thus improve data quality. In Internet surveys this visual format could provide an alternative to the more standard percent chance format. The percent chance format has been shown in previous work to yield coherent subjective distributions, but it may sound repetitive to respondents who are asked the same question about a series of different thresholds. Moreover, it requires them to remember their answers in order to respect basic properties of probabilities and to be more proficient with probabilities than would be required with the visual format.

When comparing the two formats, we show that they yield similar response rates, survey time and precision of information. In addition, the dispersion of the elicited distribution correlates with other sources of uncertainty in the expected direction for both designs suggesting that both yield high quality data from a substantive point of view.

However, the two formats differ importantly in two dimensions. First, the visual format attracts more probability in the central bin than the percent chance format. Additional experiments lead us to conclude that this is not due to anchoring towards the middle in the bins-and-balls format. We suspect that the explanation lies rather in the tendency to answer 50 percent in the percent chance questions combined with the fact that in our percent chance design respondents were not asked directly about the middle bin. In addition, we have shown how “unfolding” bins can be used to improve the precision of answers of respondents who provide very concentrated distributions in the initial bins-and-balls screen.

Second and most importantly, 20 percent of the observations are lost in the percent chance format because respondents fail to respect a basic property of probability. Moreover, those 20 percent tend to be less educated and less comfortable with probabilities than the rest of the sample introducing selection in the remaining usable answers. While this proportion may be reduced if respondents were informed about this violation, little is known about whether adding warnings or teaching respondents about probabilities improves the resulting data quality. Overall, our results favor the use of the bins-and-balls format for eliciting individuals’ subjective distribution of beliefs in Internet surveys.

Appendix 1: Construction of the mean and variance for the distribution of future Social Security benefits

We use respondents' answers about the subjective distribution of future Social Security benefits to compute individual-specific mean and variance for this distribution based on two assumptions. First, we assume that the probability mass reported by a respondent is uniformly distributed within an interval (or bin). Second, we make assumptions on the support of the distribution. For the bins-and-balls format, we assume that the extreme right bin is bounded and has the size of the other bins. For the percent chance format, we also assume that the upper interval is bounded to the right such that the size of the extreme right interval is comparable to that of the bins-and-balls format.

Based on these assumptions, we compute the expectation of the distribution as follows:

$$E(X) = \int xf(x)dx = \sum_{i=1}^{M} \int_{T_i}^{T_{i+1}} xf(x)dx = \sum_{i=1}^{M} \int_{T_i}^{T_{i+1}} xf(x)dx \,.$$

Where T_i is i's bin's left threshold, and *M=5* for the percent chance format, and *M=7* for the bins and balls format.

Under the uniform assumption, $f(x) = \frac{b_i}{\left(T_{i+1} - T_i\right)}$ for $T_{i+1} \le x \le T_i$, where b_i is the probability mass allocated between $\left[T_{i+1}, T_i\right]$.

Therefore:

$$E(X) = \sum_{i=1}^{M} \int_{T_i}^{T_{i+1}} \frac{b_i x}{\left(T_{i+1} - T_i\right)} dx = \sum_{i=1}^{M} \left[\frac{b_i x^2}{2\left(T_{i+1} - T_i\right)} \right]_{T_i}^{T_{i+1}} = \sum_{i=1}^{M} \frac{b_i \left(T_{i+1}{}^2 - T_i^2\right)}{2\left(T_{i+1} - T_i\right)} = \sum_{i=1}^{M} \frac{b_i \left(T_{i+1} + T_i\right)}{2}.$$

Based on these assumptions, we compute the variance of the distribution as follows:

$$V(X) = \int \left(x - E(X)\right)^2 f(x)dx$$

$$= \sum_{i=1}^{M} \int_{T_i}^{T_{i+1}} \frac{b_i \left(x - E(X)\right)^2}{\left(T_{i+1} - T_i\right)} dx = \sum_{i=1}^{M} \int_{T_i}^{T_{i+1}} \frac{b_i \left(x - 2xE(X) + E^2(X)\right)}{\left(T_{i+1} - T_i\right)} dx$$

$$= \sum_{i=1}^{M} \frac{b_i}{\left(T_{i+1} - T_i\right)} \left[\frac{T_{i+1}{}^3 - T_i^3}{3} - \left(T_{i+1}{}^2 - T_i^2\right) E(X) + E^2(X)\left(T_{i+1} - T_i\right) \right]$$

$$= \sum_{i=1}^{M} b_i \left[\frac{\left(T_{i+1}T_i + T_{i+1}{}^2 + T_i^2\right)}{3} - \left(T_{i+1} + T_i\right) E(X) + E^2(X) \right].$$

Appendix 2: Thresholds definition for the unfolding bins

Let L1 be the left bound of the unique bin or one of the adjacent bins, and L2 be the right bound of the unique bin or one of the adjacent bins. The following algorithm summarizes the design of the unfolding bins:

If L2-L1>=300 and L2-L1<450, use 2 bins of equal width starting at L1.

If L2-L1>=450 and L2-L1<600, use 3 bins of equal width starting at L1.

If L2-L1>=600 and L2-L1<750, use 4 bins of equal width starting at L1.

If L2-L1>=750, use 5 bins of equal width starting at L1.

So for example, if a respondent first allocated 20 balls into a bin ranging from $1,710 to $2,280, he is then asked to allocate 20 balls into 3 bins: $1,710 to $1,900, $1,900 to $2,090 and $2,090 to $2,280.

If the respondent allocated 20 balls into one bin of width smaller than $300, or into two bins of total width smaller than $450, the unfolding bins are not asked.

References

Bruine de Bruin, Wändi, Fischbeck, Paul S., Stiber, Neil A. and Fischhoff, Baruch. 2002. "What number is 'fifty-fifty'? Redistributing excess 50% responses in risk perception studies." *Risk Analysis* 22:725-735.

Bruine de Bruin, Wändi, Fischhoff, Baruch, Millstein, Susan G., and Halpern-Felsher, Bonnie L. 2000. "Verbal and Numerical Expressions of Probability: 'It's a Fifty-Fifty Chance' " Organizational Behavior and Human Decision Processes 81:115-131.

Dominitz, Jeff. 1998. "Earnings Expectations, Revisions, and Realizations." Review of Economics and Statistics, 80:374-388.

Dominitz, Jeff, and Manski Charles F. 2007. "Expected Equity Returns and Portfolio Choice: Evidence from the Health and Retirement Study." *Journal of the European Economic Association* 5:369-379.

Dominitz, Jeff, and Manski Charles F. 2006. "Measuring Pension-Benefit Expectations Probabilistically." *Labour* 20: 201-236.

Dominitz, Jeff, and Manski Charles F. 2004. "How Should We Measure Consumer Confidence?" *Journal of Economic Perspectives* 18:51-66.

Dominitz, Jeff, and Manski Charles F. 1997. "Using Expectations Data to Study Subjective Income Expectations," *Journal of the American Statistical Association*, 92:855-867.

Delavande, Adeline. Forthcoming. "Pill, Patch or Shot? Subjective Expectations and Birth Control Choice." *International Economic Review*.

Delavande, Adeline, and Rohwedder, Susann. 2007. "Eliciting Subjective Probabilities about Social Security Expectations." Rand Corporation, Working Paper.

Delavande Adeline, and Willis, Robert J. 2008. "Managing the Risk of Life," Michigan Retirement Research Center Working Paper 2007-167.

Fischhoff, Baruch, Parker, Andrew M., Bruine de Bruin, Wändi , Downs, Julie, Palmgren, Claire, Dawes, Robyn and Manski, Charles F. 2000. "Teen Expectations for Significant Life Events." *Public Opinion Quarterly* 64:89-205.

Gan, Li, Gong, Guan, Hurd, Michael D., and McFadden, Daniel L. 2004. "Subjective Mortality Risk and Bequests." NBER Working Paper No. W10789.

Hill, Daniel, Perry, Michael, Willis, Robert J. 2006. "Estimating Knightian Uncertainty from Survival Probability Questions on the HRS," University of Michigan, Working Paper.

Hurd, Michael D., and McGarry, Kathleen. 2002. "The Predictive Validity of the Subjective Probabilities of Survival." *Economic Journal* 112: 966-985.

Hurd, Michael D., and McGarry, Kathleen. 1995. "Evaluation of the Subjective Probabilities of Survival in the Health and Retirement Study." *Journal of Human Resources* 30: 268-292.

Lillard, Lee A., and Willis, Robert J. 2002. "Cognition and Wealth: The Importance of Probabilistic Thinking." Michigan Retirement Research Center Research Paper No. WP 2001-007.

Manski, Charles F. 2004. "Measuring Expectations" *Econometrica* 72:1329-1376.

Manski, Charles F., and Molinari, Francesca. 2008 "Rounding Probabilistic Expectations in Surveys." Northwestern University, Working Paper.

Table 1: Descriptive statistics of the analytical sample

	Analytical sample		HRS 2006 unweighted	HRS 2006 weighted
Variable	Frequency	Percent	Percent	Percent
Gender				
Male	535	37.8	40.2	49.0
Female	882	62.2	59.8	51.0
All	*1,417*	*100.0*	*100.0*	*100.0*
Age				
<40	11	0.8	0.7	0.0
40-44	25	1.8	1.9	0.0
45-49	90	6.4	5.9	0.0
50-54	344	24.3	26.2	25.5
55-59	660	46.5	39.1	48.1
60-64	269	19.0	22.0	24.1
65-72	19	1.3	4.3	2.3
All	*1,418*	*100.0*	*100.0*	*100.0*
Education				
Less than high school	28	2.0	12.0	9.6
High school and GED	355	25.1	30.7	29.7
some college	445	31.5	27.7	28.2
college and above	587	41.5	29.7	32.6
All	*1,415*	*100.0*	*100.0*	*100.0*

Table 2: Distribution of total number of bins used by respondents (in percent)

Number of bins	Bins and balls	Percent Chance
1	25.8	12.5
2	47.3	20.0
3	20.4	25.7
4	4.5	25.5
5	1.2	16.4
6	0.1	
7	0.8	
N	761	506

Table 3: Distribution of answers for the percent chance format (in percent)

Probability that Social Security benefits is:	Less than X1	Less than X2	More than X3	More than X4
P=0	49.3	24.1	18.2	39.7
0<P<=25	33.0	28.2	26.0	36.4
25<P<50	5.3	9.4	10.2	7.8
P=50	5.9	21.7	27.7	10.7
50<P<=75	2.6	6.9	9.9	4.2
75<P<100	2.7	6.7	6.2	0.5
P=100	1.3	3.0	1.9	0.8
N	627	627	628	627

Table 4: Distribution of answers by format

	First bin	Second bin	Middle bin	Fourth bin	Fifth bin
Bins and balls					
10th percentile	0	0	15	0	0
25th percentile	0	0	45	0	0
median	0	0	60	0	0
75th percentile	0	30	90	30	0
90th percentile	0	50	100	50	10
mean	2.74	16.98	59.93	16.96	3.38
N	761	761	761	761	761
Percent chance					
10th percentile	0	0	0	0	0
25th percentile	0	0	15	0	0
median	0	10	40	15	5
75th percentile	10	25	70	25	20
90th percentile	30	49	100	50	50
mean	9.57	16.13	43.02	18.17	13.12
N	506	506	506	506	506

Table 5: Distribution of the point expectations and computed means of the elicited distribution

	Point expectation	Computed Mean (Bins and balls)	Computed Mean (Percent chance)
10th percentile	500	497.0	468.6
25th percentile	750	715.0	753.3
median	1100	1040.0	1111.5
75th percentile	1500	1417.0	1559.3
90th percentile	1859	1837.5	1976.5
mean	1157.9	1127.1	1186.3
N	1,363	751	506

Table 6: Distribution of answers excluding respondents allocating all the probability mass in the central bin

	Percent of respondents
Bins and Balls (N=565)	
More probability mass to the right of the central bin	49.5
More probability mass to the left of the central bin	45.5
Equal probability on both sides of the central bin	5.5
Percent (N=443)	
More probability mass to the right of the central bin	47.8
More probability mass to the left of the central bin	31.6
Equal probability on both sides of the central bin	20.5

Table 7: Distribution of balls for the expected claiming age

	61 or less	62	63	64	65	66	67	68 or more
10th percentile	0	0	0	0	0	0	0	0
25th percentile	0	0	0	0	0	0	0	0
median	0	0	0	0	2	0	0	0
75th percentile	0	6	2	2	8	4	2	0
90th percentile	0	16	5	5	15	10	8	10
mean	0.35	4.03	1.52	1.65	4.81	2.89	2.04	2.27
N	1,017	1,017	1,017	1,017	1,017	1,017	1,017	1,017

Table 8: Distribution of expected Social Security amount in the ALP

Social Security expected amount	10th perc.	25th perc.	median	75th perc.	90th perc.	mean	N
Snowball sample - 8th wave	200	500	1000	1400	2266	1086.20	73
Snowball sample - 24th wave	300	600	857	1200	1900	1002.46	95
All respondents aged 30-64 - 8th wave	400	750	1200	1500	2000	1206.08	650
Snowball aged 30-64 - 8th wave (weighted)	400	600	900	1300	2000	1067.92	95

Table 9: Distribution of balls for the ALP snowball sample: waves 8 versus wave 24

	First bin	Second bin	Third bin	Fourth bin	Fifth bin	Sixth bin	Seventh bin
8th wave				(contains point estimate)			
10th percentile	0	0	0	0	0	0	0
25th percentile	0	0	0	6	0	0	0
median	0	0	2	9	3	0	0
75th percentile	0	1	6	12	7	1	0
90th percentile	1	5	10	18	9	3	1
mean	1.04	1.40	3.32	9.47	3.44	0.95	0.40
N	73	73	73	73	73	73	73
24th wave			(contains point estimate)				
10th percentile	0	0	2	0	0	0	0
25th percentile	0	0	6	0	0	0	0
median	0	2	10	3	0	0	0
75th percentile	0	7	13	7	1	0	0
90th percentile	4	10	18	10	4	1	0
mean	1.27	3.68	9.44	3.98	0.94	0.47	0.21
N	95	95	95	95	95	95	95

Table 10: Distribution of balls for the ALP: wave 8 versus weighted wave 24

	First bin	Second bin	Third bin	Fourth bin	Fifth bin	Sixth bin	Seventh bin
8th wave (all respondents aged 30-64)				(contains point estimate)			
10th percentile	0	0	0	4	0	0	0
25th percentile	0	0	0	8	0	0	0
median	0	0	2	10	3	0	0
75th percentile	0	0	6	14	6	0	0
90th percentile	0	2	10	20	10	2	0
mean	0.43	0.57	3.57	10.81	3.70	0.68	0.23
N	650	650	650	650	650	650	650
24th wave (all respondents aged 30-64) weighted			(contains point estimate)				
10th percentile	0	0	3	0	0	0	0
25th percentile	0	0	8	0	0	0	0
median	0	2	10	3	0	0	0
75th percentile	0	7	14	7	0	0	0
90th percentile	4	10	20	10	3	1	0
mean	1.17	3.64	10.23	3.70	0.77	0.39	0.10
N	95	95	95	95	95	95	95

Table 11: Number of unfolding bins administered

Number of bins	Freq.	Percent
2	40	21.5
3	59	31.7
4	33	17.7
5	54	29.0
Total	186	100.0

Table 12: Number of unfolding bins used by respondents

Number of bins presented	Number of bins used			Total
	1	2	3	
2	80.6	19.4		100.0
3	29.1	65.5	5.5	100.0
4	21.9	68.8	9.4	100.0
5	9.8	66.7	23.5	100.0
Total	32.8	56.9	10.3	100.0

Table 13: Distributions of balls by unfolding bins

Two unfolding bins					
	First bin	Second Bin			
median	0	20			
mean	4.67	15.33			
N	36	36			
Three unfolding bins					
	First bin	Second Bin	Third Bin		
median	0	13	0		
mean	2.64	12.87	4.49		
N	55	55	55		
Four unfolding bins					
	First bin	Second Bin	Third Bin	Fourth Bin	
median	0	10	10	0	
mean	2.00	7.69	8.44	1.88	
N	32	32	32	32	
Five unfolding bins					
	First bin	Second Bin	Third Bin	Fourth Bin	Fifth Bin
median	0	2	8	6	0
mean	1.16	5.39	6.96	5.98	0.51
N	51	51	51	51	51

Table 14: Average number of bins and probability mass in the middle bin by characteristics

	Average number of bins		Average probability mass in the middle bin	
	Bins and balls	Percent chance	Bins and balls	Percent chance
Gender				
male	2.08	3.11	61.97	44.23
female	2.07	3.15	59.36	42.24
Number of bins used for expected claiming age				
1	1.62***	2.67 ***	63.98*	50.05**
2	1.89	3.17	60.83	45.11
3	2.25	3.30	59.09	37.75
4	2.45	3.37	51.67	36.48
5+	2.90	3.55	51.45	40.06
Distance to expected claiming age				
5 years of less	1.81***	2.66***	65.47***	55.20***

6 to 10 years	2.05	3.21	61.59	41.25
more than 10 years	2.26	3.52	55.62	33.56
Subjective probability of receiving SS in the future				
P<95%	2.25***	3.43***	55.33***	34.26**
P>=95%	1.91	2.87	64.75	50.73
Subjective probability of reform that would reduce own Social Security benefits within 10 years				
P<50%	1.89***	2.81***	65.40***	55.31***
P>=50%	2.14	3.29	58.37	37.09
Subjective probability of working past age 62				
P<=20%	1.97**	3.10	62.89	43.31
20%<P<=80%	2.19	3.23	58.66	41.82
P>80%	2.10	3.20	56.95	42.93
Subjective probability of loosing job during the next year				
P=0%	1.92***	2.91***	60.19	42.34
0%<P<=10%	2.14	3.48	61.07	44.70
P>10%	2.18	3.21	63.25	39.80
Self-reported health				
excellent/very good	2.13**	3.14	58.39*	41.55*
good	1.98	3.17	63.48	49.47
poor/fair	1.89	3.00	64.22	39.00
Expect to claim on own Social Security record only				
No	2.18**	3.15	57.75*	44.81
Yes	2.03	3.16	62.22	43.87

* indicates that, using a F-test, we can reject that all averages are equal at 10%, ** at 5% and *** at 1%

Table 15: Multivariate analysis

Dependent variable	Number of bins used		Probability in the central bin		Imputed standard deviation	
Model	Ordered probit		OLS		OLS	
Random design	Bins and balls	Percent chance	Bins and balls	Percent chance	Bins and balls	Percent chance
Female	-0.063	0.058	-1.729	-1.298	-52.711***	-64.909***
	[0.479]	[0.596]	[0.465]	[0.676]	[0.000]	[0.004]
Married	-0.013	0.026	0.639	4.635	2.421	-14.474
	[0.907]	[0.817]	[0.821]	[0.153]	[0.834]	[0.537]
Some college	0.060	0.169	1.483	-0.369	15.057	50.088**
	[0.529]	[0.139]	[0.555]	[0.910]	[0.142]	[0.035]
Years to expected claiming age	0.025***	0.029***	-0.299	-1.095***	2.885***	12.352***
	[0.002]	[0.006]	[0.166]	[0.000]	[0.001]	[0.000]
Total number of bins used for claiming age	0.328***	0.132***	-1.874*	-2.280**	19.040***	15.497**
	[0.042]	[0.036]	[1.094]	[0.998]	[4.469]	[7.227]
Expect to receive Social Security on own record only	-0.121	0.090	2.242	-2.250	16.219	47.805*
	[0.104]	[0.124]	[2.754]	[3.562]	[11.247]	[25.806]
Excellent/very good health	-	-	-	-	-	-
Good	-0.123	0.098	3.803	7.815**	-9.161	-13.809
	[0.097]	[0.123]	[2.553]	[3.486]	[10.429]	[25.249]
Poor/fair	-0.291*	-0.072	7.526*	-0.832	-28.152*	-57.724
	[0.157]	[0.179]	[4.067]	[5.108]	[16.611]	[37.006]
Work for pay	-0.986	-1.702**	20.312	-32.201	-12.134	-211.561
	[0.855]	[0.836]	[21.038]	[22.243]	[85.929]	[161.127]
Has pension on current job	-0.076	0.059	7.777***	-1.776	30.857***	28.803
	[0.103]	[0.120]	[2.731]	[3.438]	[11.153]	[24.904]
Subjective probability of loosing job during the next year	0.006***	-0.001	0.041	-0.003	0.266	-1.284**
	[0.002]	[0.003]	[0.063]	[0.074]	[0.257]	[0.536]
Subjective probability of working past age 62	0.000	0.000	-0.057*	0.040	0.372***	0.153
	[0.001]	[0.001]	[0.034]	[0.042]	[0.137]	[0.301]
Subjective probability of reform that would reduce own Social Security benefits within 10 years	0.004***	0.007***	-0.094**	-0.233***	-0.001	0.998***
	[0.001]	[0.002]	[0.037]	[0.050]	[0.151]	[0.365]
Subjective probability of receiving SS in the future	-0.005**	-0.004	0.144***	0.128*	0.136	0.737
	[0.002]	[0.002]	[0.055]	[0.071]	[0.224]	[0.513]
Lowest wealth tercile	-	-	-	-	-	-
Middle wealth tercile	0.043	-0.076	-1.482	2.543	-2.787	-1.369
	[0.118]	[0.139]	[3.099]	[3.987]	[12.656]	[28.879]

Highest wealth tercile	0.120	-0.096	-2.783	11.361***	17.086	-41.948
	[0.118]	[0.145]	[3.118]	[4.140]	[12.734]	[29.987]
Constant			35.956	83.307***	74.004	290.476
			[22.742]	[24.742]	[92.891]	[179.236]
Observations	*742*	*506*	*742*	*506*	*742*	*506*
R-squared			0.09	0.19	0.16	0.19
Pseudo R-squared	0.08	0.05				

standard errors in brackets

* significant at 10%; ** significant at 5%; *** significant at 1%

All the regressions include dummy for missing values (coefficients not shown)

Figure 1: Bins-and-balls format

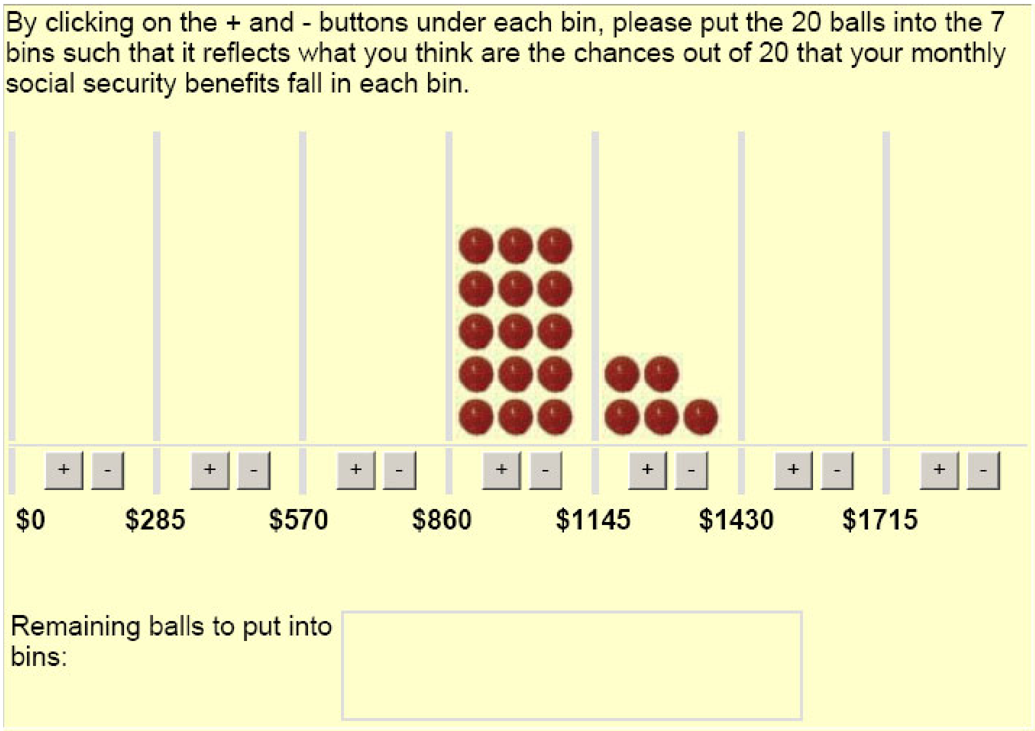
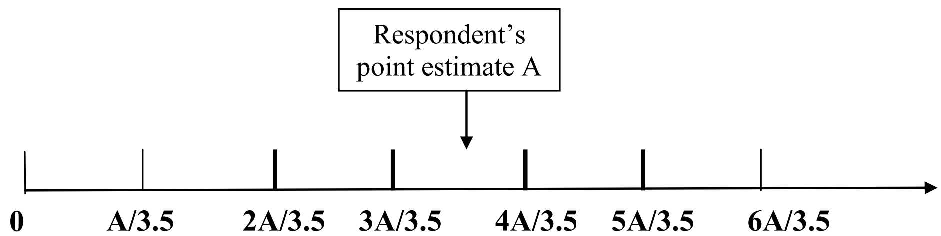


Figure 2: Threshold design

Visual format



Percent chance format

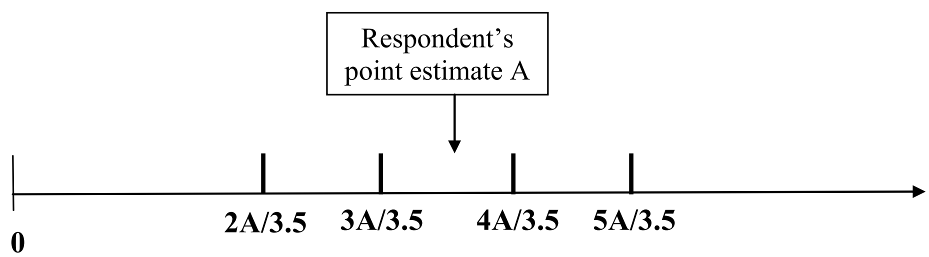
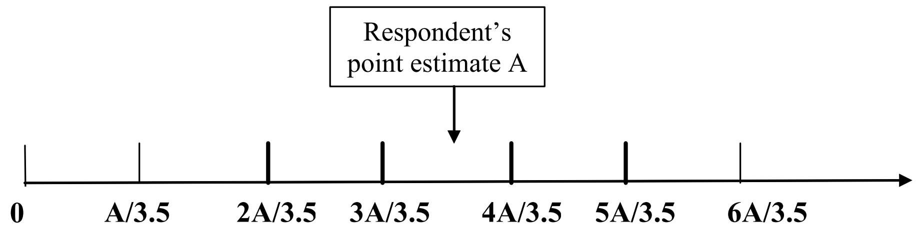


Figure 3: Thresholds definition in HRS Internet and in the ALP

Visual format in HRS Internet and 8th wave of the ALP



Visual format in the 24th wave of the ALP