On Probabilistic Expectations and Rounding in Surveys

Fabian Gouret

On Probabilistic Expectations and Rounding in Surveys *

Fabian Gouret[†]

February 16, 2017

Abstract

Answers to subjective probability questions in surveys are mainly heaped to different degrees at multiples of 5 or 10, suggesting that respondents round their responses. Taking these responses at face value is clearly problematic for gross form of rounding, given that a respondent gives a point response when only a wide interval response is meaningful. This paper analyzes three subjective probability questions posed in the Survey of Economic Expectations. The specific format of these questions permits, under some mild assumptions, to develop for each question a measure related to the extent of rounding. Whatever the question, a substantial fraction of answers are rounded grossly; it includes a majority of the 50s, as well as numerous M10. We are also able to find that younger respondents are more likely to provide a gross rounding to one of the question which asks them their expectations concerning the Social Security benefits that they will collect when they will turn 70.

JEL: C81, C83, D8

Keywords: probabilistic expectations, rounding, interval data, partial identification.

[†]Théma, Université de Cergy-Pontoise, 33 Bvd du Port, 95011 Cergy-Pontoise Cedex, France (Email: fabiangouret@gmail.com). Homepage: https://sites.google.com/site/fabgouret/

1 Introduction

Applied econometricians have increasingly measured subjective probabilistic expectations that individuals hold about the future realization of various economic variables. Subjective probabilistic expectations permit to study how beliefs can affect some risky behaviors, and improve the empirical content of structural models of decision-making under uncertainty.[1] However, answers to subjective probability questions are mainly heaped to different degrees at multiples of 5 or 10, suggesting that respondents round their responses.[2] The problem remains relatively innocuous when respondents round their responses to the nearest 1, 5 or perhaps 10 percent. But taking probabilistic responses at face value is clearly problematic for grosser form of rounding, given that a respondent gives a point response when only a wide interval response would make sense. The grossest form of rounding occurs when one respondent provides a point response but he is in fact unable to assign any subjective probability in the interval $[0, 100]$. This raises various questions: Is gross rounding a problem (i.e., is the estimated probability that a response is gross rounding relatively substantial and statistically significant)? What are the responses which are rounded grossly? Why sample respondents give probabilistic expectations which are rounded grossly?

Most of the literature on probabilistic expectations has taken subjective probabilistic responses at face value. The few exceptions which consider the possibility of rounding usually impose restrictions on the extent of rounding that each possible answer may reflect. de Bresser and van Soest (2013) and Kleinjans and van Soest (2014) develop a model with endogenous regime switching and unobserved regimes to identify the extent of rounding. They assume that only some 50s may reflect the maximal interval [0,100], or a very large interval [25,75] with a width of 50 percentage points.[3] They assume that

---

[1]For stochastic models of choice under uncertainty, an important contribution is Delavande (2008). She combines data on expectations for the effectiveness of different contraceptive methods and choices to estimate a random utility model (RUM) of contraception behavior. Delavande and Kohler (forthcoming) who propose a RUM to study how risky sexual behavior is influenced by individuals' expectations about own and partner's HIV status also propose in their introduction an extensive review of the models of choice under uncertainty that have been estimated so far.

[2]As in Manski and Molinari (2010, p.219), I mean by rounding "*the familiar practice of reporting one value whenever a real number lies in an interval*".

[3]Otherwise, a 50 may reflect the intervals [37.5,62.5], [45,55], [47.5,52.5] or [49.5,50.5]. Some distributional assumptions permit then to achieve point-identification of the quantities of interest.

the multiples of ten other than $(0, 50, 100)$ –M10– (i.e., 10, 20, 30, 40, 60, 70, 80 and 90 percent) are in the range [M10-5,M10+5] in the worst case[4], i.e., the width of the interval cannot exceed 10 percentage points. The multiples of 5 but not of ten –M5– except 25 and 75 percent (i.e., 5, 15, 35,...65, 85, 95 percent) are assumed to imply the interval [M5-2.5,M5+2.5] in the worst case. Manski and Molinari (2010, p.223) use each person's response pattern across a class of expectations questions in the Health and Retirement Study (HRS) to infer the extent to which the person rounds responses to particular questions. If at least one of the responses to a class of questions is a M10 or a M5, they assume that his responses imply maximal intervals of 10 and 5 percentage points respectively. Manski and Molinari (2010) rightly explain that "*research interpreting reported expectations as interval data makes weaker assumptions than does research taking responses at face value*" (p.220) but "*the actual rounding interval may not be a subset of the [assumed] interval*" (p.223). In such a case, the use of "*the data is not correct*" (p.223).

The present paper builds upon Gouret (2015) (Gouret hereafter) who develops a method that I relate here to the extent of rounding. This method, called a measure of coherence or imprecision, is applicable on a specific format of probabilistic questions used to know the distribution of a continuous variable, but it is based on less restrictions than the methods cited above; this format was introduced initially in the Survey of Economic Expectations (SEE). I extend the method and the analysis in various ways. First, while Gouret uses it to filter the data and conserves at the end only the responses which are relatively precise, I extend the method to provide a measure of rounding and estimate the probability that a response is gross rounding, the probability that a M10 is gross rounding and so on; this permits, in turn, to test statistically various hypotheses made in the literature. Second, contrary to Gouret who only focuses on a question of the SEE which measures the probabilistic beliefs that Americans hold about the stock market in the year ahead, the current paper also analyzes two other important continuous variables that were elicited in the SEE: the probabilistic beliefs that Americans

---

[4]Otherwise, a M10 may reflect the intervals [M10-2.5,M10+2.5] or [M10-0.5,M10+0.5]. Again, some distributional assumptions permit to achieve point-identification of the quantities of interest.

(under 70) hold concerning the Social Security benefits that they will collect when they will turn 70, and the probabilistic beliefs that Americans hold about their income in the year ahead. Lastly, I try to understand if the probability to provide a gross rounding to these different questions is related to some covariates.

Let me describe the design of these three questions. Each question j follows a two-stage questioning method. The first stage is composed of *two preliminary questions* which ask each respondent i for $r_{i,j,min}$ and $r_{i,j,max}$, the lowest and highest possible values that the continuous variable $R_{i,j}$ might take in a future date. The second asks for a *sequence of K probabilistic questions* of the type: "*What is the percent chance that $R_{i,j}$ would be worth over $r_{i,j,k}$?*" The sequence of specific thresholds $r_{i,j,1} < r_{i,j,2} < \ldots < r_{i,j,K}$ is chosen among a finite number of possible sequences by an algorithm that uses the average of the responses $r_{i,j,min}$ and $r_{i,j,max}$ to the first-stage questions.[5] The probabilistic responses to the second stage are denoted $Q_{i,j,k} \equiv \mathrm{P}\left(R_{i,j} > r_{i,j,k}\right)$, $k = 1, 2, \ldots, K$.[6] As an example, Table 1 provides the stock market question ($j = S$). Tables A1 and A2 in Appendix A provides the Social Security benefits question ($j = B$) and the income question ($j = I$). Remark that the sequence of probabilistic responses has to decrease given the monotonicity of the complementary cumulative distribution function ($Q_{i,j,1} \geq Q_{i,j,2} \geq \ldots \geq Q_{i,j,K}$). Given that $Q_{i,j,1}$ is the sole probabilistic answer which does not depend on any other probabilistic answers, I will focus on the rounding of $Q_{i,j,1}$.

Section 2 explains how this two-stage questioning method permits to test the appropriateness of some practices in the literature. I first build upon Gouret to propose a measure which is related to the extent of rounding. The method is based on the common idea that a respondent engages in gross rounding when he has a very imprecise belief on an event (because of a lack of knowledge, or because he has only a finite amount of time to process information and provide an answer). A respondent who has a precise subjective distribution in mind concerning $R_{i,j}$ should use the same underlying subjective

---

[5]The two preliminary questions permit to have an idea of the support of the distribution. Additional reasons for asking these preliminary questions can be found in Manski (2004, footnote 17, pp.1346-1347).

[6]Instead of asking K points on the respondent i's subjective complementary cumulative distribution function of $R_{i,I}$, the income question ($j = I$) asks points on the respondent i's subjective cumulative distribution; see Subsection A.2 in Appendix A. One thus observes the probabilistic answers $(100 - Q_{i,j,k}) \equiv [100 - \mathrm{P}\left(R_{i,j} > r_{i,j,k}\right)]$, $k = 1, 2, \ldots, K$. It does not change however the analysis.

SEE scenario

    The next question is about investing in the stock market. Please think about the type of mutual fund known as a diversified stock fund. This type of mutual fund holds stock in many different companies engaged in a wide variety of business activities. Suppose that tomorrow someone were to invest one thousand dollars in such a mutual fund. Please think about how much money this investment would be worth one year from now.

Preliminary questions

    What do you think is the LOWEST amount that this investment of \$1000 would possibly be worth one year from now? ($r_{i,S,min}$)

    What do you think is the HIGHEST amount that this investment of \$1000 would possibly be worth one year from now? ($r_{i,S,max}$)

Algorithm for selection of investment thresholds

$\left\lceil \frac{r_{i,S,min}+r_{i,S,max}}{2} \right\rceil$	$r_{i,S,1}$	$r_{i,S,2}$	$r_{i,S,3}$	$r_{i,S,4}$
0 to 899	500	900	1000	1100
900 to 999	800	900	1000	1100
1000 to 1099	900	1000	1100	1200
1100 to 1299	1000	1100	1200	1500
1300 or more	1000	1200	1500	2000

Note: The midpoint of the lowest and highest values rounded up to the next integer $\left(\left\lceil \frac{r_{i,S,min}+r_{i,S,max}}{2} \right\rceil \right)$ determines the four investment value thresholds $r_{i,S,k}$, $k = 1, 2, 3, 4$, according to the algorithm presented in this Table.

Sequence of $K = 4$ probabilistic questions

    What do you think is the PERCENT CHANCE (or CHANCES OUT OF 100) that, one year from now, this investment would be worth over $r_{i,S,k}$? ($Q_{i,S,k}$, $k = 1, 2, 3, 4$)

distribution to answer the two stages of question j. The possible incoherence between the responses $r_{i,j,min}$ and $r_{i,j,max}$ to the first stage, i.e., the lowest and highest possible values of $R_{i,j}$, and the first subjective probability $Q_{i,j,1}$ provided to the second stage is exploited to have an idea of the extent of rounding. Note that if one imposed too strong assumptions to have a precise measure of rounding, their price would be the loss of credibility of the conclusions. The approach developed is based consciously on a mild assumption which says that the interval $(r_{i,j,min}, r_{i,j,max})$ is not even the support stricto sensu of the distribution, but only a suggestive support, as considered by Dominitz and Manski (1997). This simple assumption permits to learn imperfectly something about the extent of rounding of some answers because the algorithms which select the thresholds do not insure that all of them belong to $(r_{j,min}, r_{j,max})$. The first threshold $r_{j,1}$ is in practice often outside this interval, in particular lower than $r_{j,min}$ (between 72 and 89 percent of the respondents depending on the question j); these algorithms are described in Table 1 for the stock market question and in Tables A1 and A2 (in Appendix A)

for the Social Security and the income questions. Under this mild assumption, the different relative frequencies studied are partially identified (the relative frequency that a response (to question j) is gross rounding, the relative frequency that a 50 is gross rounding, the relative frequency that a M10 is gross rounding...), but it is enough to test various hypotheses made in the literature; see Manski (2003) for an introduction to partial identification.

Section 3 analyzes the responses to the three questions of the SEE and shows that the relative frequency that a response is gross rounding is relatively high and significantly different from zero whatever the question. I also find that the relative frequency that a 50 is rounded grossly is always high (above 50 percent) and statistically significant. Finally, the relative frequency that a M10 is a gross rounding is important and always statistically significant.

Section 4 studies if the probability to provide a gross rounding is related to some covariates. Given that the probability to provide a gross rounding is partially identified, very few results emerge. Notwithstanding, we are able to find a result which makes sense: compared to older respondents, younger respondents are more likely to provide a gross rounding to the Social Security benefit question. The concluding Section 5 suggests directions for future research.

2   Testing the appropriateness of some practices in the literature

Subsection 2.1 relates the extent of rounding to imprecise subjective probabilities. It builds upon Gouret and describes how the specific format of questions used in the SEE can be useful to measure the degree of imprecision in the respondent's answers concerning an event, a measure related to the extent of rounding. Subsection 2.2 considers weak assumptions to provide an interval measurement of imprecision. Subsection 2.3 explains how this measure can be used to learn about different probabilities and test various hypotheses made in the literature.

2.1 Imprecise subjective probabilities and rounding

Consider a survey respondent i who has responded to a question j which concerns the future realization of a continuous variable $R_{i,j}$. This question j follows the two-stage questioning method of the SEE described in the Introduction: the first stage asks for $r_{i,j,min}$ and $r_{i,j,max}$, the lowest and highest possible values that $R_{i,j}$ might take; the second stage elicits a sequence of K probabilities $Q_{i,j,k} \equiv \mathrm{P}\left(R_{i,j} > r_{i,j,k}\right)$, $k = 1, 2, \ldots, K$, where the thresholds $r_{i,j,1} < r_{i,j,2} < \ldots < r_{i,j,K}$ are set by using the responses to the first stage. As already noticed, I focus on the rounding of $Q_{i,j,1}$ given that this subjective probability determines the subjective probabilities for the subsequent thresholds.

To understand why the two-stage questioning method permits to learn something about the extent of rounding of $Q_{i,j,1}$, note that the literature commonly says that a respondent engages in rounding when he does not hold a unique subjective distribution for a future event but a set of subjective distributions. For instance, a survey respondent may round because he perceives *"the future as partially ambiguous and, hence, not feel able to place precise probabilities"* on an event (Manski and Molinari, 2010, p.220). A respondent may also have enough information to have a precise probability in principle, but processing information may be cognitively costly and a survey respondent has only a finite amount of time to provide an answer.[7] If a respondent has a precise subjective distribution in mind concerning $R_{i,j}$, he should use the same underlying subjective distribution to answer the two stages of question j. The idea is thus to see if each respondent uses the same underlying subjective distributions for $R_{i,j}$. A first subjective distribution is used by the respondent to report $r_{i,j,min}$ and $r_{i,j,max}$, the lowest and highest possible values of this first subjective distribution. Let $\widetilde{\mathrm{P}}\left(R_{i,j} > r\right)$ be the subjective probability of the event $R_{i,j} > r$ associated to this first distribution. A second subjective distribution is used by the respondent to answer the sequence of probabilistic questions $Q_{i,j,k} \equiv \mathrm{P}\left(R_{i,j} > r_{i,j,k}\right)$, $k = 1, 2, \ldots, K$.

Suppose for the moment that a researcher has full knowledge of the first subjective distribution

---

[7] As pointed out by Manski and Molinari (2010, p.220), another reason is that *"some persons may hold precise subjective probabilities for future events, as presumed in Bayesian statistics, but round their responses to simplify communication."*

used by the respondent to answer $r_{i,j,min}$ and $r_{i,j,max}$. Then, he can assess the subjective probability of the event $R_{i,j} > r_{i,j,1}$ according to this first distribution, i.e., $\widetilde{P}(R_{i,j} > r_{i,j,1})$. Subjective probabilities are imprecise when $\widetilde{P}(R_{i,j} > r_{i,j,1})$ and $Q_{i,j,1}$ disagree. A researcher can conclude that the respondent believes that the percent chance of the event $R_{i,j} > r_{i,j,1}$ is *at least* in the range $\left[Q_{i,j,1}, \widetilde{P}(R_{i,j} > r_{i,j,1})\right]$ if $Q_{i,j,1} \leq \widetilde{P}(R_{i,j} > r_{i,j,1})$; or at least in the range $\left[\widetilde{P}(R_{i,j} > r_{i,j,1}), Q_{i,j,1}\right]$ if $Q_{i,j,1} \geq \widetilde{P}(R_{i,j} > r_{i,j,1})$. For instance, if $Q_{i,j,1} = 95$ but $\widetilde{P}(R_{i,j} > r_{i,j,1}) = 100$, one can conclude that the respondent believes that the percent chance of the event $R_{i,j} > r_{i,j,1}$ is *at least* in the range $[95, 100]$; so $Q_{i,j,1}$ may be a small rounding. If $Q_{i,j,1} = 60$ but $\widetilde{P}(R_{i,j} > r_{i,j,1}) = 90$, one can conclude that the respondent believes that the percent chance of the event $R_{i,j} > r_{i,j,1}$ is at least in the range $[60, 90]$; so $Q_{i,j,1}$ is clearly a gross rounding.

The width of the interval, or, equivalently, the degree of imprecision concerning the event $R_{i,j} > r_{i,j,1}$ is then:

$$d_{i,j,1} \equiv \left| Q_{i,j,1} - \widetilde{P}(R_{i,j} > r_{i,j,1}) \right| \tag{1}$$

$d_{i,j,1} \in [0, 100]$ is called a measure of coherence in Gouret. He considers that when $d_{i,j,1}$ is too high, i.e., when $Q_{i,j,1}$ and $\widetilde{P}(R_{i,j} > r_{i,j,1})$ are clearly incoherent, $Q_{i,j,1}$ conveys no information, but he does not exploit the value of $d_{i,j,1}$ per se. Here, I prefer to use the phrase "degree of imprecision" because the value $d_{i,j,1}$ has a clear meaning: it reflects the degree of imprecision in percentage points in the respondent i's answers concerning the event $R_{i,j} > r_{i,j,1}$. If $Q_{i,j,1} = 60$ but $\widetilde{P}(R_{i,j} > r_{i,j,1}) = 90$, one can conclude that the respondent believes that the percent chance of the event $R_{i,j} > r_{i,j,1}$ is at least in the range $[60, 90]$, and $d_{i,j,1} = 30$ tells us that the width of this interval is 30 percentage points (or, equivalently, that the degree of imprecision is 30 percentage points).

It is important to understand that even if $\widetilde{P}(R_{i,j} > r_{i,j,1})$ were observed, the interval considered for respondent i would be a subset of the true interval, and $d_{i,j,1}$ would underestimate the true degree of imprecision. For instance, if a respondent reports $Q_{i,j,1} = 50$ because he feels unable to assign any subjective probability in the interval $[0, 100]$, then the true interval for respondent i is $[0, 100]$ and the

true degree of imprecision is 100 percentage points, as the theoretical literature on imprecise subjective probabilities would point out (e.g., Walley, 1991, p.210). However, given that $Q_{i,j,1} = 50$, the maximal value that $d_{i,j,1}$ can take is $d_{i,j,1} = 50$ percentage points; it occurs when $\widetilde{P}(R_{i,j} > r_{i,j,1}) = 100$ or 0. Nevertheless, if one knows that $d_{i,j,1} = 50$, it is enough to say at least that $Q_{i,j,1} = 50$ reflects a gross rounding.

2.2   An interval measurement of imprecision

The difficulty is that the subjective probability at the first threshold according to the first subjective distribution $\widetilde{P}(R_{i,j} > r_{i,j,1})$ is unobserved. In fact, a researcher knows little about the first subjective distribution, except that $r_{i,j,min}$ and $r_{i,j,max}$ are the lowest and highest possible values of the first subjective distribution. Furthermore, Dominitz and Manski (1997) note that the answers $r_{i,j,min}$ and $r_{i,j,max}$ are not interpretable literally as minimum and maximum values. This claim may be (i.) difficult to understand, but it is (ii.) impossible to disagree with them. First, this claim may be (i.) difficult to understand. Analyzing responses to an household income question which asks points on the respondent i's subjective cumulative distribution (like the income question in Appendix A.2), Dominitz and Manski (1997, p.860) write precisely:

> "$[r_{j,min} \in (r_{j,1}, \ldots, r_{j,4})]$ for 200 of the 437 respondents and $[r_{j,max} \in (r_{j,1}, \ldots, r_{j,4})]$ for 221 respondents. Among the 200 respondents asked for the percent chance that household income will be less than $[r_{j,min}]$, the median response was 20%. Among the 221 respondents asked for the percent chance that income will be less than $[r_{j,max}]$ the median response was 80%. [...] These findings indicate that most respondents associate the phrases "lowest possible" and "highest possible" with low and high probabilities, but do not interpret these phrases as defining the support of a subjective distribution."

The last sentence remains an interpretation. They could have written equally that these discrepancies reflect the difficulty for the respondents to have a precise subjective probability, as the incoherence

seems to suggest. Indeed, It remains strange that a respondent who is unable to provide precisely the support of his distribution will then be able to provide precise subjective probabilities.[8] It is however (ii.) difficult to disagree with Dominitz and Manski in the sense that saying that the responses $r_{i,j,min}$ and $r_{i,j,max}$ to the two preliminary questions only suggest the support of the distribution is a weaker assumption than saying that $r_{i,j,min}$ and $r_{i,j,max}$ provide the support stricto sensu. Instead of assuming $\widetilde{P}(R_{i,j} \leq r_{i,j,min}) = 0$ and $\widetilde{P}(R_{i,j} \geq r_{i,j,max}) = 0$, I will thus make the following weaker assumption:

Assumption 1. $0 \leq \widetilde{P}(R_{i,j} \leq r_{i,j,min}) \leq \alpha$ and $0 \leq \widetilde{P}(R_{i,j} \geq r_{i,j,max}) \leq \alpha$

Assumption 1 states the $\widetilde{P}(R_{i,j} \leq r_{i,j,min})$ and $\widetilde{P}(R_{i,j} \geq r_{i,j,max})$ can be any value in the interval $[0, \alpha]$. In other words, the econometrician does not know if the respondent i has interpreted the phrases "lowest possible" and "highest possible" literally as minimum and maximum values (i.e., $\widetilde{P}(R_{i,j} \leq r_{i,j,min}) = 0$ and $\widetilde{P}(R_{i,j} \geq r_{i,j,max}) = 0$), or if he has associated these phrases with low and high probabilities (i.e., $\widetilde{P}(R_{i,j} \leq r_{i,j,min}) \in (0, \alpha]$ and $\widetilde{P}(R_{i,j} \geq r_{i,j,max}) \in (0, \alpha]$). The higher α, the weaker the assumption. However, if one can argue that some respondents may associate "lowest possible" and "highest possible" with "low" and "high" values, nobody will say that these phrases are so vague that some respondents may interpret them as "medium" values. If so, the upper bound α in Assumption 1 should remain low. I will consider $\alpha = 20$ in the empirical analysis. I do not think that any researcher will consider that it is not enough. And all researchers who would have considered a lower value for α must accept $\alpha = 20$. For instance, if one researcher assumes that 10 is enough, i.e., $0 \leq \widetilde{P}(R_{i,j} \leq r_{i,j,min}) \leq 10$, he must accept $0 \leq \widetilde{P}(R_{i,j} \leq r_{i,j,min}) \leq 20$; the same for $\widetilde{P}(R_{i,j} \geq r_{i,j,max})$.

Under Assumption 1, the degree of imprecision $d_{i,j,1}$ is interval measured. That is, $d_{i,j,1}$ is not observed but rather we observe $d_{i,j,1,L}$ and $d_{i,j,1,U}$, such that $d_{i,j,1,L} \leq d_{i,j,1} \leq d_{i,j,1,U}$ for all i if a researcher accepts $\alpha = 20$. In other words, the researcher has an imprecise knowledge on the degree

---

[8]Note that the elicitation of a sequence of quantiles instead of a sequence of probabilities has been common in decision analysis to construct subjective probability distributions for continuous variables (see, e.g., Lichtenstein *et al.*, 1982). Recently, Abbas *et al.* (2008) have compared in an online experiment the two methods. Most of the participants express a clear preference for the elicitation of probabilities in the postexperimental evaluations, but the two methods provide very similar results. However the distributions of the continuous variable extracted from the elicitation of probabilities "*fit the historical data slightly better*" (p.197). This small superiority of the elicitation of a sequence of probabilities is in particular due to "*the nature of the response scale [...], which is bounded by 0 (impossibility) and 1 (certainty)*" contrary to the quantiles.

of imprecision $d_{i,j,1}$ that the respondent i faces concerning the event $R_{i,j} > r_{i,j,1}$. To understand, consider first that $r_{i,j,1} \leq r_{i,j,min}$. If this is the case, Assumption 1 permits to say that $100 - \alpha \leq \widetilde{P}(R_{i,j} > r_{i,j,1}|r_{i,j,1} \leq r_{i,j,min}) \leq 100$. It is then easy to find that if $r_{i,j,1} \leq r_{i,j,min}$:

$$d_{i,j,1,L} = \max\{0, 100 - Q_{i,j,1} - \alpha\} \leq d_{i,j,1} \leq \max\{100 - Q_{i,j,1}, |100 - Q_{i,j,1} - \alpha|\} = d_{i,j,1,U} \quad (2)$$

Now if $r_{i,j,1} \geq r_{i,j,max}$, then $0 \leq \widetilde{P}(R_{i,j} > r_{i,j,1}|r_{i,j,1} \geq r_{i,j,max})) \leq \alpha$, and:

$$d_{i,j,1,L} = \max\{0, Q_{i,j,1} - \alpha\} \leq d_{i,j,1} \leq \max\{Q_{i,j,1}, |Q_{i,j,1} - \alpha|\} = d_{i,j,1,U} \quad (3)$$

Lastly, if $r_{i,j,1} \in (r_{i,j,min}, r_{i,j,max})$, then $0 \leq \widetilde{P}(R_{i,j} > r_{i,j,1}|r_{i,j,1} \in (r_{i,j,min}, r_{i,j,max})) \leq 100$, and:

$$d_{i,j,1,L} = 0 \leq d_{i,j,1} \leq \max\{Q_{i,j,1}, 100 - Q_{i,j,1}\} = d_{i,j,1,U} \quad (4)$$

It is important to distinguish between the different intervals which appear in our analysis. There is the set of subjective probabilities on the event $R_{i,j} > r_{i,j,1}$ that a respondent i may have in mind and which lies at least in the interval $\left[Q_{i,j,1}, \widetilde{P}(R_{i,j} > r_{i,j,1})\right]$ (if $Q_{i,j,1} \leq \widetilde{P}(R_{i,j} > r_{i,j,1})$). The width of this interval, i.e., the degree of imprecision $d_{i,j,1}$ (Equation 1), may be a function of the degree of information that the respondent has on the event $R_{i,j} > r_{i,j,1}$, as well as his cognitive capacity to process information. However, because the researcher has an imperfect knowledge on $\widetilde{P}(R_{i,j} > r_{i,j,1})$, the degree of imprecision is interval measured, i.e., $d_{i,j,1}$ belongs to the interval $[d_{i,j,1,L}, d_{i,j,1,U}]$. $d_{i,j,1}$ is interval measured because Assumption 1 is weak. Adding assumptions would permit to reduce the interval $[d_{i,j,1,L}, d_{i,j,1,U}]$. However, the econometric analysis with additional assumptions that cannot be justified by substantive arguments may generate strong but unreliable conclusions, as often highlighted (e.g., Manski, 2003; Schollmeyer and Augustin, 2013, pp.2-3). Thus, we will not add additional assumptions.

2.3 Relative frequencies and hypothesis testing

Consider now a survey of N respondents drawn at random from the population of interest. For each $i = 1, \ldots, N$, we have an interval measurement of $d_{i,j,1}$ because the interval $[d_{i,j,1,L}, d_{i,j,1,U}]$ is observable but $d_{i,j,1}$ is not.

The first important question is: is gross rounding a problem? Most of the literature on probabilistic expectations takes the responses at face value, and thus considers that there is no gross rounding (nor small rounding). If so, one would like to learn the probability that a response has a degree of imprecision of more than $\underline{d}$ percentage points, i.e., $P(d_{j,1} \geq \underline{d})$, and test the null hypothesis $H_0 : P(d_{j,1} \geq \underline{d}) = 0$. If $\underline{d} \in (0, 100)$ is sufficiently high and the null is rejected whatever the question j, then it will cast serious doubts on the papers which take probabilistic responses at face value and do not consider the possibility of gross rounding. If $d_{i,j,1}$ were observed, the natural estimator of $P(d_{j,1} \geq \underline{d})$ would be the relative frequency of responses whose degree of imprecision is more than $\underline{d}$, i.e., $P_N(d_{j,1} \geq \underline{d}) = \frac{\sum_{i=1}^{N} \mathbf{1}[d_{i,j,1} \geq \underline{d}]}{N}$. In the absence of additional assumptions other than Assumption 1 on the precise value of $d_{i,j,1}$ in the interval $[d_{i,j,1,L}, d_{i,j,1,U}]$, the relative frequency $P_N(d_{j,1} \geq \underline{d})$ is partially identified. Interval measurement of the variables $d_{j,1}$ yields very simple lower and upper bounds on $P_N(d_{j,1} \geq \underline{d})$: the relative frequency $P_N(d_{j,1,L} \geq \underline{d}) = \frac{\sum_{i=1}^{N} \mathbf{1}[d_{i,j,1,L} \geq \underline{d}]}{N}$ is the lower bound on $P_N(d_{j,1} \geq \underline{d})$, while $P_N(d_{j,1,U} \geq \underline{d}) = \frac{\sum_{i=1}^{N} \mathbf{1}[d_{i,j,1,U} \geq \underline{d}]}{N}$ is the upper bound on $P_N(d_{j,1} \geq \underline{d})$ (see, e.g., Manski, 2003, p.18, Proposition 1.4). Note that $P_N(d_{j,1,U} \geq \underline{d}) = P_N(d_{j,1,L} \geq \underline{d}) + P_N(d_{j,1,U} \geq \underline{d} > d_{j,1,L})$, with $P_N(d_{j,1,U} \geq \underline{d} > d_{j,1,L}) = \frac{\sum_{i=1}^{N} \mathbf{1}[d_{i,j,1,U} \geq \underline{d} > d_{i,j,1,L}]}{N}$; such decomposition will be useful for the ease of exposition. The identification region for $P_N(d_{j,1} \geq \underline{d})$ is then:

$$H\left[P_N(d_{j,1} \geq \underline{d})\right] = \left[P_N(d_{j,1,L} \geq \underline{d}), P_N(d_{j,1,L} \geq \underline{d}) + P_N(d_{j,1,U} \geq \underline{d} > d_{j,1,L})\right] \tag{5}$$

where $H[\cdot]$ denotes the identification region for the quantity in brackets. Logically, the width of this interval is the relative frequency of responses for which it is impossible to say if their degree of

imprecision is higher or lower than $\underline{d}$ percentage points, i.e., $P_N(d_{j,1,U} \geq \underline{d} > d_{j,1,L})$.

To take into consideration sampling variation and test statistically the null hypothesis $H_0 : P(d_{j,1} \geq \underline{d}) = 0$, I will follow a confidence-interval approach. I will use the Bonferroni inequality to obtain an asymptotic confidence interval which covers the identification region $H[P(d_{j,1} \geq \underline{d})]$ with probability at least 99 percent; this confidence interval necessarily covers $P(d_{j,1} \geq \underline{d})$ with probability at least 99 percent, given that $P(d_{j,1} \geq \underline{d})$ lies in $H[P(d_{j,1} \geq \underline{d})]$; see Appendix B for more details. If zero is always outside the estimated confidence intervals, the null is always rejected.

A second important question concerns the unease that researchers have felt about the high prevalence of 50s, suspecting that a substantial share of them reflects maximally or at least very imprecise subjective probabilities (Bruine de Bruin *et al.*, 2000; Manski, 2004, p.1370; Hudomiet and Willis, 2013). If so, one would like to learn the relative frequency of 50s whose degree of imprecision is more than $\underline{d}$ percentage points, i.e., $P_N(d_{j,1} \geq \underline{d}|Q_{j,1} = 50)$, and test the null $H_0 : P(d_{j,1} \geq \underline{d}|Q_{j,1} = 50) = 0$. Again, one should choose a $\underline{d}$ sufficiently high. And, again, because $d_{i,j,1}$ is interval measured, $P_N(d_{j,1} \geq \underline{d}|Q_{j,1} = 50)$ is partially identified, and its identification region is $H[P_N(d_{j,1} \geq \underline{d}|Q_{j,1} = 50)]$. I will also provide a Bonferroni confidence interval which covers this identification region with probability at least 99 percent to see if the null is rejected or not.

The last but not the least important question concerns the few authors (Manski and Molinari, 2010, Kleinjans and van Soest, 2014) who interpret subjective probabilities as interval data but consider that the answers other than (0,50,100) cannot imply intervals of more than 10 percentage points (except 25 and 75 in the case of Kleinjans and van Soest, 2014). For instance, one may want to learn the relative frequency of M10 whose degree of imprecision is more than $\underline{d}$ percentage points, i.e., $P_N(d_{j,1} \geq \underline{d}|Q_{j,1} = M10)$, and test the null $H_0 : P(d_{j,1} \geq \underline{d}|Q_{j,1} = M10) = 0$. As before, $P_N(d_{j,1} \geq \underline{d}|Q_{j,1} = M10)$ is partially identified because $d_{i,j,1}$ is interval measured. If $\underline{d}$ is higher than 10 percentage points and the null is rejected, it will cast serious doubt on the intervals assumed by the authors mentioned above.

Before to step any further, I have to define a plausible value for $\underline{d}$. Obviously I have to choose a value higher than 10 percentage points. The higher $\underline{d}$, the more conservative will be the approach. I will focus on $\underline{d} = 50 - \alpha$. Various reasons explain this choice. First, all readers will agree that this critical value is extremely conservative: if $\alpha = 20$, then choosing $\underline{d} = 30$ means that a response is considered as gross rounding if the degree of imprecision concerning the event $R_{i,j} > r_{i,j,1}$ is at least 30 percentage points. Second, consider the unease of the researchers concerning the heap of 50s in the surveys, suspecting that some of them reflect gross rounding (or even the grossest form of rounding). If $Q_{i,j,1} = 50$ and $r_{i,j,1} \notin (r_{i,j,min}, r_{i,j,max})$, the interval measurement of $d_{i,j,1}$ is $[50 - \alpha, 50]$ (see Equations 2 and 3), while it is $[0,50]$ when $Q_{j,1} = 50$ and $r_{i,j,1} \in (r_{i,j,min}, r_{i,j,max})$ (see Equation 4). We are unable to see if some of them reflect the grossest form of rounding because the maximal value that the degree of imprecision can take is 50 percent points. But if $Q_{i,j,1} = 50$ and $r_{i,j,1} \notin (r_{i,j,min}, r_{i,j,max})$, we can say that the degree of imprecision is at least $50 - \alpha$ percentage points, so at least 30 percentage points if $\alpha = 20$. Third, as explained in Propositions 1 and 2, the identification region for $\mathrm{P}_N(d_{j,1} \geq \underline{d}|Q_{j,1} = 50)$ is similar for all values of $\underline{d}$ and α which satisfy $\underline{d} \leq 50 - \alpha$, and the lower bound on $\mathrm{P}_N(d_{j,1} \geq \underline{d})$ is similar for all values of $\underline{d}$ and α which satisfy $\underline{d} = 50 - \alpha$. These properties of invariance may be seen as useful.

Proposition 1. Consider that for each respondent $i = 1, \ldots, N$, $d_{i,j,1}$ is interval measured, i.e., we observe the interval $[d_{i,j,1,L}, d_{i,j,1,U}]$ as given by Equations 2, 3 and 4. Then the identification region for $\mathrm{P}_N(d_{j,1} \geq \underline{d}|Q_{j,1} = 50)$ is similar for all combinations of $\underline{d}$ and α which satisfy $\underline{d} \leq 50 - \alpha$ and is given by:

$$
\begin{aligned}
\mathrm{H}\left[\mathrm{P}_N(d_{j,1} \geq \underline{d}|Q_{j,1} = 50)\right] \quad = \quad & [\mathrm{P}_N(d_{j,1,L} \geq \underline{d}|Q_{j,1} = 50), \\
& \mathrm{P}_N(d_{j,1,L} \geq \underline{d}|Q_{j,1} = 50) + \mathrm{P}_N(d_{j,1,U} \geq \underline{d} > d_{j,1,L}|Q_{j,1} = 50)]
\end{aligned}
\tag{6}
$$

where:

$$P_N(d_{j,1,L} \geq \underline{d}|Q_{j,1} = 50) = \frac{\sum\limits_{i=1}^{N} \mathbf{1}\left[r_{i,j,1} \notin (r_{i,j,min}, r_{i,j,max}), Q_{i,j,1} = 50\right]}{\sum\limits_{i=1}^{N} \mathbf{1}\left[Q_{i,j,1} = 50\right]} \text{ , and}$$

$$P_N(d_{j,1,U} \geq \underline{d} > d_{j,1,L}|Q_{j,1} = 50) = \frac{\sum\limits_{i=1}^{N} \mathbf{1}\left[r_{i,j,1} \in (r_{i,j,min}, r_{i,j,max}), Q_{i,j,1} = 50\right]}{\sum\limits_{i=1}^{N} \mathbf{1}\left[Q_{i,j,1} = 50\right]}$$

Proof: see Appendix C.1.

Proposition 2. Consider that for each respondent $i = 1, \ldots, N$, $d_{i,j,1}$ is interval measured, i.e., we observe the interval $[d_{i,j,1,L}, d_{i,j,1,U}]$ as given by Equations 2, 3 and 4. If $\underline{d} = 50 - \alpha$, then the identification region for $P_N(d_{j,1} \geq \underline{d})$ is:

$$H\left[P_N(d_{j,1} \geq 50 - \alpha)\right] = \left[P_N(d_{j,1,L} \geq 50 - \alpha), P_N(d_{j,1,L} \geq 50 - \alpha) + P_N(d_{j,1,U} \geq 50 - \alpha > d_{j,1,L})\right]$$

where $P_N(d_{j,1,L} \geq 50 - \alpha) = \dfrac{\sum\limits_{i=1}^{N} \mathbf{1}\left[d_{i,j,1,L} \geq 50 - \alpha\right]}{N}$ is identical for all combinations of $\underline{d}$ and α which satisfy $\underline{d} = 50 - \alpha$.

Proof: see Appendix C.2.

Proposition 1 provides the lower and upper bounds on the relative frequency of 50s which have a degree of imprecision higher than $\underline{d}$ for all $\underline{d}$ and α which satisfy $\underline{d} \leq 50 - \alpha$. If $r_{i,j,1} \notin (r_{i,j,min}, r_{i,j,max})$ and $Q_{i,j,1} = 50$, the interval measurement of $d_{i,j,1}$ is $[50 - \alpha, 50]$ as already noticed (see Equations 2 and 3); so $d_{i,j,1,L} = 50 - \alpha \geq \underline{d}$ for all $\underline{d} \leq 50 - \alpha$ if $r_{i,j,1} \notin (r_{i,j,min}, r_{i,j,max})$ and $Q_{i,j,1} = 50$. But if $r_{i,j,1} \in (r_{i,j,min}, r_{i,j,max})$ and $Q_{i,j,1} = 50$, the interval measurement of $d_{i,j,1}$ is $[0, 50]$ (see Equation 4). So $d_{i,j,1,L} = 0 < \underline{d}$ for all $\underline{d} \leq 50 - \alpha$ if $r_{i,j,1} \in (r_{i,j,min}, r_{i,j,max})$. Hence, the lower bound on $P_N(d_{j,1} \geq \underline{d}|Q_{j,1} = 50)$ is the relative frequency of 50s which are provided to a threshold outside the suggestive support, i.e., $P_N(d_{j,1,L} \geq \underline{d}|Q_{j,1} = 50) = \frac{\sum_{i=1}^{N} \mathbf{1}[r_{i,j,1} \notin (r_{i,j,min}, r_{i,j,max}), Q_{i,j,1} = 50]}{\sum_{i=1}^{N} \mathbf{1}[Q_{i,j,1} = 50]}$ for all $\underline{d} \leq 50 - \alpha$. Con-

14

cerning the upper bound, note that $d_{i,j,1,U} = 50$ for all i. So the upper bound on $P_N(d_{j,1} \geq \underline{d}|Q_{j,1} = 50)$

is $P_N(d_{j,1,U} \geq \underline{d}|Q_{j,1} = 50) = \frac{\sum_{i=1}^{N} \mathbf{1}[r_{i,j,1} \notin (r_{i,j,min}, r_{i,j,max}), Q_{i,j,1}=50]}{\sum_{i=1}^{N} \mathbf{1}[Q_{i,j,1}=50]} + \frac{\sum_{i=1}^{N} \mathbf{1}[r_{i,j,1} \in (r_{i,j,min}, r_{i,j,max}), Q_{i,j,1}=50]}{\sum_{i=1}^{N} \mathbf{1}[Q_{i,j,1}=50]}$

for all $\underline{d} \leq 50 - \alpha$.

Proposition 2 states that the lower bound on the identification region for $P_N(d_{j,1} \geq \underline{d})$ is the same for all combinations of α and $\underline{d}$ which satisfy $\underline{d} = 50 - \alpha$. That is, if $\{\alpha, \underline{d}\} = \{20, 30\}$ or if $\{\alpha, \underline{d}\} = \{0, 50\}$, then $P_N(d_{j,1,L} \geq 50 - \alpha)$ is the same. One can naturally ask: if the lower bound on the relative frequency of responses which are gross rounding is the same for all combinations of α and $\underline{d}$ which satisfy $\underline{d} = 50 - \alpha$, why should I choose $\{\alpha, \underline{d}\} = \{20, 30\}$, and not $\{\alpha, \underline{d}\} = \{0, 50\}$? Again, $P_N(d_{j,1,L} \geq \underline{d})$ is only the lower bound on $P_N(d_{j,1} \geq \underline{d})$, the fraction of responses which are gross rounding. The upper bound $P_N(d_{j,1,U} \geq \underline{d})$ on $P_N(d_{j,1} \geq \underline{d})$ is lower if $\{\alpha, \underline{d}\} = \{0, 50\}$ than if $\{\alpha, \underline{d}\} = \{20, 30\}$. To understand, note that the relative frequency of responses for which it is impossible to say if their degree of imprecision is higher or lower than $\underline{d}$ percentage points, i.e., $P_N(d_{j,1,U} \geq \underline{d} > d_{j,1,L})$, is lower if $\{\alpha, \underline{d}\} = \{0, 50\}$: if $\alpha = 0$, the degree of imprecision would be no more interval measured when $r_{i,j,1} \notin (r_{i,j,min}, r_{i,j,max})$ because $d_{i,j,1,L} = d_{i,j,1,U}$ (see Equations 2 and 3); if so, a researcher would always know precisely if the degree of imprecision is higher or lower than $\underline{d}$ percentage points. Thus, the weaker assumption $\alpha = 20$ increases the upper bound on $P_N(d_{j,1} \geq \underline{d})$ (see Equation 5).

3   Empirical analysis

This Section uses the method developed in Section 2 to analyze the responses to the three expectations questions of the SEE: the stock market question ($j = S$), the income question ($j = I$) and the Social Security benefit question ($j = B$). Subsection 3.1 describes the survey and the data. Subsection 3.2 provides the results.

3.1 The data

The SEE was a periodic module in WISCON, a continuous national computer-assisted telephone survey conducted by the University of Wisconsin Survey Center; Dominitz and Manski (2004) describe the basic features of the survey. The SEE module which elicited expectations of significant personal events was included in 15 waves over the period April 1994 to May 2002.[9] Income, stock market and Social Security benefits expectations were the three main continuous variables that were elicited.[10] Contrary to the income question which was elicited in all waves, the Social Security question was only elicited in waves 12, 13, 14 and 15, and the stock market question in wave 12, 13 and 14. This paper thus focuses on the three waves in which all the three questions were included: wave 12 (where interviews were conducted in the period July 1999-November 1999), wave 13 (February 2000-May 2000) and wave 14 (September 2000-March 2001).

1651 respondents were interviewed in these three cross-sections (547 in wave 12, 465 in wave 13, and 639 in wave 14). The stock market and income questions were posed to these 1651 respondents (of age 18 or older). But logically the Social Security question, which elicited the beliefs that respondents hold concerning the Social Security benefits that they will collect when they will be 70 years old, concerned only the persons of age 18-69, i.e., 1425 respondents in the survey. Furthermore, before the Social Security question, i.e., the two preliminary questions and the sequence of probabilistic questions described in Table A1, the respondents were requested to predict their eligibility for benefits. Only the respondents of age 18-69 who provided a non-zero probability were asked the preliminary questions and the sequence of probabilistic questions; see Appendix A for more details. Of the 1425 respondents, 1280 provided a positive probability of eligibility, so were asked the preliminary questions and the sequence of probabilistic questions (406 in wave 12, 376 in wave 13, and 498 in wave 14).

---

[9]After the continuous survey ceased, an additional wave of interviews, wave 16, was conducted from a special "omnibus" survey in the period October to November 2002.

[10]In the 11 first waves, another continuous variable was elicited: the respondents looking for work had to provide their subjective cumulative distribution function about the time to find a job. I do not consider this question for two reasons: first and essentially, less than 200 respondents were concerned by this question (so less than 20 respondents per wave on average); second, and as explained below, this paper focuses on waves 12, 13 and 14.

Some of these respondents did not answer $Q_{j,1}$, in particular because of the skip-sequencing feature of the SEE: those who did not answer the preliminary questions were not asked the sequence of probabilistic questions. At the stock market question, of the 1651 respondents, 1255 answered the preliminary questions $r_{S,min}$ and $r_{S,max}$. Of these 1255 respondents, 1231 provided an answer $Q_{S,1}$ (415 in wave 12, 342 in wave 13, and 474 in wave 14). We can compute the bounds of the measure of imprecision at the first threshold for these 1231 respondents, so for 75 percent of the respondents $\left(\simeq \frac{1231}{1651}\right)$. Concerning the income question, 1260 respondents answered the preliminary questions $r_{I,min}$ and $r_{I,max}$; of these 1260 respondents, 1247 provided an answer $Q_{I,1}$ (409 in wave 12, 357 in wave 13, and 481 in wave 14), so the effective response rate is 76 percent $\left(\simeq \frac{1247}{1651}\right)$ and very similar to the one of the stock market question. Lastly, of the 1280 respondents concerned by the Social Security question, only 860 answered the preliminary questions $r_{B,min}$ and $r_{B,max}$; of these 860 respondents, 851 provided an answer $Q_{B,1}$ (253 in wave 12, 244 in wave 13, and 354 in wave 14). The effective response rate of 66 percent $\left(\simeq \frac{851}{1280}\right)$ is much lower than those at the two other questions.

This Section tries to understand if some respondents who provided a response $Q_{i,j,1}$ engaged in gross rounding and the patterns of responses. It is not interested in nonresponses; see Section 4 for that. However, it is interesting to note that $r_{i,j,min}$ and $r_{i,j,max}$ are missing at important rates. This is particularly true for the Social Security question. Most of these nonresponses are some "don't know" and very few refusals. One can interpret these "don't know" as a response strategy for respondents who have highly imprecise beliefs concerning $R_{i,j}$: perhaps they lack any relevant information or thinking is so costly that they do not want to put any effort in their answers. Thus, instead of providing very gross rounding, they say that they "don't know". In line with this view, Appendix D shows that nonresponses are mainly related to education whatever the question j (given that one can reasonably assume that the cognitive cost to provide a precise subjective probability is more important for less educated respondents); it also shows that at the Social Security question, younger respondents are more likely to provide a "don't know" than older respondents (obviously, a respondent who is 20 years

17

old has probably highly imprecise beliefs concerning the Social Security benefits that he will collect when he will be 70 years old).[11]

Lastly, remember that when $r_{i,j,1} \in (r_{i,j,min}, r_{i,j,max})$, we know nothing about $\widetilde{P}(R_{i,j} > r_{i,j,1})$, the subjective probability at the first threshold according to the first subjective distribution, so $d_{i,j,1,L} = 0$ and $d_{i,j,1,U} \geq 50$ whatever the response $Q_{i,j,1}$ (see Equation 4). In such case, we do not know if the respondent has a precise belief on the event $R_{i,j} > r_{i,j,1}$, or a very imprecise one. If $r_{i,j,1} \in (r_{i,j,min}, r_{i,j,max})$ for all i, then the lower bounds on $P_N(d_{j,1} \geq \underline{d})$, $P_N(d_{j,1} \geq \underline{d}|Q_{j,1} = 50)$, $P_N(d_{j,1} \geq \underline{d}|Q_{j,1} = M10)$... must be zero (because $d_{i,j,1,L} = 0 < \underline{d} = 50 - \alpha$ whatever the response $Q_{i,j,1}$), and the upper bounds must be 1 (because $d_{i,j,1,U} \geq 50 > \underline{d} = 50 - \alpha$ whatever the response $Q_{i,j,1}$). In other words, the identification regions of the relative frequencies of interest will be [0,1] and we will learn nothing. It is thus important that $r_{j,1} \notin (r_{j,min}, r_{j,max})$ for some respondents. It happens that $r_{S,1} \notin (r_{S,min}, r_{S,max})$ for 878 of the 1231 respondents at the stock market question, so 72 percent (more precisely, $r_{S,1} \leq r_{S,min}$ for 70 percent, and $r_{S,1} \geq r_{S,max}$ for 2 percent); $r_{I,1} \notin (r_{I,min}, r_{I,max})$ for 1070 of the 1247 respondents at the income question, so 86 percent (more precisely, $r_{I,1} \leq r_{I,min}$ for 77 percent, and $r_{I,1} \geq r_{I,max}$ for 9 percent); $r_{B,1} \notin (r_{B,min}, r_{B,max})$ for 753 of the 851 respondents at the Social Security question, so 89 percent (more precisely, $r_{B,1} \leq r_{B,min}$ for 70 percent, and $r_{B,1} \geq r_{B,max}$ for 19 percent).

3.2   Results

Table 2 provides the raw data. The column labeled $Q_{j,1} = 50$ gives the number of respondents who answered "50 percent". The column labeled $Q_{j,1} = 100$ gives the number of respondents who answered

---

[11]This interpretation of "don't know" is in line with the literature on the psychology of survey response (e.g., Tourangeau *et al.*, 2000, pp.250-254) which highlights that saying "don't know", like answering randomly or selecting the same response at various thresholds, are response strategies for a boundedly rational respondent who do not consider the different stages involved by answering a question (i.e., understanding the question, retrieving relevant information, using this information to make a judgment, and selecting an answer). Tourangeau *et al.* consider that a respondent can use a weak or strong satisficing criterion. In their terminology, weak satisficing occurs when the respondent is less thoughtful about the different stages that a survey question involves and picks the first response which seems adequate rather than the one that is optimal because the cognitive costs entailed by making an optimal choice becomes high; strong satisficing occurs when the respondent omits whole stages, and uses some response strategies like selecting "don't know", choosing the middle of the scale, or answering randomly.

"100 percent". The column labeled $Q_{j,1} = 0$ gives the number of respondents who answered "0 percent". The column labeled $Q_{j,1} = 75$ gives the number of respondents who answered "75 percent", while the one labeled $Q_{j,1} = 25$ gives the number of those who answered "25 percent". Column $Q_{j,1} =$M10 gives the number of respondents who answered a multiple of ten other than (0,50,100): e.g., 20, 30 or 90 percent. Column $Q_{j,1} =$M5 gives the number of respondents who answered a multiple of 5 but not of ten, other than 25 and 75 percent: e.g., 5,15 or 85 percent. The column labeled $Q_{j,1} =$"Other" gives the number of respondents who provided an answer which is not a multiple of 5: e.g., 99, 98 or 1 percent. For each column, the raw data are provided by question, wave, and category. By category, I mean those whose degree of imprecision is clearly below $\underline{d} = 30$ percentage points (i.e., $\underline{d} > d_{i,j,1,U}$), those whose degree of imprecision is clearly above 30 percentage points (i.e., $d_{i,j,1,L} \geq \underline{d}$), ie., those who are clearly gross rounding, and those for which it is impossible to say if their degree of imprecision is higher or lower than 30 percentage points (i.e., $d_{i,j,1,U} \geq \underline{d} > d_{j,1,L}$). Such categorization is useful. It permits to compute very easily the estimated bounds for the relative frequencies of interest.

Consider first the relative frequency of responses which are gross rounding. In wave 12, of the 409 respondents who provided a subjective probability $Q_{i,I,1}$ to the income question, $d_{I,1,L} \geq \underline{d}$ for 52 respondents, so $P_N(d_{I,1,L} \geq \underline{d}|$ Wave=12$) = \frac{52}{409} \simeq 0.127$; and $d_{I,1,U} \geq \underline{d} > d_{I,1,L}$ for 82, so $P_N(d_{I,1,U} \geq \underline{d}|$ Wave=12$) = \frac{52+82}{409} \simeq 0.327$. The resulting estimate of the bound on $P_N(d_{I,1} \geq \underline{d}|$ Wave=12$)$ is $[0.127, 0.327]$. As shown in Part [A] of Table 3, similar results are found for the other waves. The relative frequency of gross rounding provided to the income question is always substantial, given that the lower bound is at least 12 percent. Table 3 also provides the result for the other questions. Results are very similar across waves in both cases. At the stock market question, the estimate of the bound on $P_N(d_{S,1} \geq \underline{d})$ is $[0.194, 0.533]$ when the three waves are pooled, while it is $[0.294, 0.485]$ for the Social Security question. Remark also that zero is always outside the Bonferroni 99 percent confidence intervals of the identification regions, so the null $H_0 : P(d_{j,1} \geq \underline{d}) = 0$ is rejected for all questions j. These first results cast serious doubts on the papers which ignore the possibility that probabilistic

Table 2: Responses tendencies by question, wave and category (based on Assumption 1 with $\alpha = 20$ and $\underline{d} = 30$)

	$Q_{j,1}=50$	$Q_{j,1}=100$	$Q_{j,1}=0$	$Q_{j,1}=75$	$Q_{j,1}=25$	$Q_{j,1}$=M10	$Q_{j,1}$=M5	$Q_{j,1}$=Other	Tot.
Stock market question ($j = S$)									
Wave 12									
$\#\{\underline{d} > d_{i,S,1,U}\}$		59	3	21	1	74	25	15	198
$\#\{d_{i,S,1,L} \geq \underline{d}\}$	56		3		3	14	6	4	86
$\#\{d_{i,S,1,U} \geq \underline{d} > d_{i,S,1,L}\}$	24	4		11	2	77	9	4	131
Total	80	63	6	32	6	165	40	23	415
Wave 13									
$\#\{\underline{d} > d_{i,S,1,U}\}$		63		19	1	55	19	8	165
$\#\{d_{i,S,1,L} \geq \underline{d}\}$	44	1	3			18	2	3	71
$\#\{d_{i,S,1,U} \geq \underline{d} > d_{i,S,1,L}\}$	22	6	2	10		55	11		106
Total	66	70	5	29	1	128	32	11	342
Wave 14									
$\#\{\underline{d} > d_{i,S,1,U}\}$		67	3	27		85	21	9	212
$\#\{d_{i,S,1,L} \geq \underline{d}\}$	50		7		3	17	3	3	83
$\#\{d_{i,S,1,U} \geq \underline{d} > d_{i,S,1,L}\}$	40	8	1	25	1	87	13	4	179
Total	90	75	11	52	4	189	37	16	474
Income question ($j = I$)									
Wave 12									
$\#\{\underline{d} > d_{i,I,1,U}\}$		15	109	3	4	79	36	29	275
$\#\{d_{i,I,1,L} \geq \underline{d}\}$	32	3	3	4		8	1	1	52
$\#\{d_{i,I,1,U} \geq \underline{d} > d_{i,I,1,L}\}$	13	4	4	2	2	44	9	4	82
Total	45	22	116	9	6	131	46	34	409
Wave 13									
$\#\{\underline{d} > d_{i,I,1,U}\}$		12	101		9	64	46	26	258
$\#\{d_{i,I,1,L} \geq \underline{d}\}$	26	3	3	2		8	1		43
$\#\{d_{i,I,1,U} \geq \underline{d} > d_{i,I,1,L}\}$	12	2	5	3	3	23	5	3	56
Total	38	17	109	5	12	95	52	29	357
Wave 14									
$\#\{\underline{d} > d_{i,I,1,U}\}$		29	155	2	8	63	44	30	331
$\#\{d_{i,I,1,L} \geq \underline{d}\}$	27	4	6	3	1	13	3	2	59
$\#\{d_{i,I,1,U} \geq \underline{d} > d_{i,I,1,L}\}$	20	4	4	3	1	47	9	3	91
Total	47	37	165	8	10	123	56	35	481
Social Security question ($j = B$)									
Wave 12									
$\#\{\underline{d} > d_{i,B,1,U}\}$		38	15	12	3	47	14	12	141
$\#\{d_{i,B,1,L} \geq \underline{d}\}$	31	4	7		2	14	2	1	61
$\#\{d_{i,B,1,U} \geq \underline{d} > d_{i,B,1,L}\}$	8		2		1	35	3	2	51
Total	39	42	24	12	6	96	19	15	253
Wave 13									
$\#\{\underline{d} > d_{i,B,1,U}\}$		41	8	9	2	37	17	4	118
$\#\{d_{i,B,1,L} \geq \underline{d}\}$	42	5	3	2	5	25	3	1	86
$\#\{d_{i,B,1,U} \geq \underline{d} > d_{i,B,1,L}\}$	7			2	4	24	3		40
Total	49	46	11	13	11	86	23	5	244
Wave 14									
$\#\{\underline{d} > d_{i,B,1,U}\}$		65	16	9	4	58	14	13	179
$\#\{d_{i,B,1,L} \geq \underline{d}\}$	51	8	7		3	28	6	1	104
$\#\{d_{i,B,1,U} \geq \underline{d} > d_{i,B,1,L}\}$	9	3	1	3		46	7	2	71
Total	60	76	24	12	7	132	27	16	354

responses may be gross rounding.

Consider now the unease that researchers eliciting probabilistic expectations have felt about the heap of 50s. As shown in Table 2, at the income question, of the 45 respondents who answered $Q_{i,I,1} = 50$ in wave 12, $d_{I,1,L} \geq \underline{d}$ for 32 respondents, so $P_N(d_{I,1,L} \geq \underline{d}|Q_{I,1} = 50, \text{Wave}=12) = \frac{32}{45} \simeq 0.711$; and $d_{I,1,U} \geq \underline{d} > d_{I,1,L}$ for 13, so $P_N(d_{I,1,U} \geq \underline{d}|Q_{I,1} = 50, \text{Wave}=12) = \frac{32+13}{45} = 1$. The resulting estimate of the bound on $P_N(d_{I,1} \geq \underline{d}|Q_{I,1} = 50, \text{Wave}=12)$ is $[0.711, 1]$. One might argue that interpretation of this estimated bound should be cautious, given the small number of 50s (45). But Table 3 shows that the results for the other waves are broadly similar, and when the three waves are pooled, $H[P_N(d_{I,1} \geq \underline{d}|Q_{I,1} = 50)] = [0.653, 1]$. Similar comments can be made for the stock market and the Social Security questions, and we do find that $H[P_N(d_{S,1} \geq \underline{d}|Q_{S,1} = 50)] = [0.636, 1]$ and $H[P_N(d_{B,1} \geq \underline{d}|Q_{B,1} = 50)] = [0.838, 1]$ when the three waves are pooled. It is also possible to test the null $H_0 : P(d_{j,1} \geq \underline{d}|Q_{j,1} = 50) = 0$, and Part (B) in Table 3 confirms that it is always rejected. All these results confirm that researchers are right to feel embarrassed with the 50s, given that an important and significant proportion of them are clearly rounded grossly whatever the wave and the question.

Lastly, the few authors (Manski and Molinari, 2010, Kleinjans and van Soest, 2014) who interpret subjective probabilities as interval data consider that the answers other than (0,50,100) cannot imply intervals of more than 10 percentage points (except 25 and 75 in the case of Kleinjans and van Soest, 2014). It is interesting to see in Table 2 that some M10, few M5 and few responses which are not multiples of 5 are clearly gross rounding whatever the question. Concerning the M5 and the responses which are not multiples of 5, note however the small number of respondents who provided these answers. For instance, in wave 12, at the income question, only 46 persons responded a M5, and $d_{I,1,L} \geq \underline{d}$ for only one of them. Even if we pool the three waves, the lower bound on the relative frequency of M5 which are gross rounding will be very close to zero ($P_N(d_{I,1,L} \geq \underline{d}|Q_{I,1} = M5) = \frac{1+1+3}{46+52+56} \simeq 0.03$). The M10 deserve more attention. In wave 12, at the Social Security question,

21

between 0.145 and 0.510 M10 are rounded grossly, as shown in Part (C) of Table 3. In waves 13 and 14, the lower bounds are even higher, given that $\text{H}\left[P_N(d_{B,1} \geq \underline{d}|Q_{B,1} = \text{M10, Wave=13})\right] = [0.291, 0.569]$ and $\text{H}\left[P_N(d_{B,1} \geq \underline{d}|Q_{B,1} = \text{M10, Wave=14})\right] = [0.212, 0.560]$. For the stock market and the income questions, the lower bounds are smaller but remain non-negligible ($\text{H}\left[P_N(d_{S,1} \geq \underline{d}|Q_{S,1} = \text{M10})\right] = [0.102, 0.556]$ and $\text{H}\left[P_N(d_{I,1} \geq \underline{d}|Q_{I,1} = \text{M10})\right] = [0.083, 0.409]$ when the three waves are pooled). These results seem to contradict the literature which usually assumes that a M10 cannot imply an interval with a width of more than 10 percentage points. If we test formally the null $H_0 : P(d_{j,1} \geq \underline{d}|Q_{j,1} = \text{M10}) = 0$, one can see that it is rejected whatever the wave and the question, given that zero is always outside the Bonferroni 99 percent confidence intervals of the identification regions.

Table 3: Estimated identification region for $P(d_{j,1} \geq \underline{d})$, $P(d_{j,1} \geq \underline{d}|Q_1 = 50)$ and $P(d_{j,1} \geq \underline{d}|Q_1 = \text{M10})$ and estimated asymptotic Bonferroni confidence intervals with level at least 99 percent ($\alpha = 20$ and $\underline{d} = 30$)

		Wave 12	Wave 13	Wave 14	Total
[A]					
	Stock market question $\text{H}\left[P_N(d_{S,1} \geq \underline{d})\right]$	[0.207,0.523] (0.156,0.586)	[0.207,0.517] (0.151,0.587)	[0.175,0.552] (0.130,0.611)	[0.194,0.533] (0.165,0.569)
	Income question $\text{H}\left[P_N(d_{I,1} \geq \underline{d})\right]$	[0.127,0.327] (0.084,0.387)	[0.120,0.277] (0.076,0.338)	[0.122,0.312] (0.084,0.366)	[0.123,0.307] (0.099,0.340)
	Social Security question $\text{H}\left[P_N(d_{B,1} \geq \underline{d})\right]$	[0.241,0.442] (0.171,0.523)	[0.352,0.516] (0.273,0.598)	[0.293,0.494] (0.231,0.563)	[0.294,0.485] (0.254,0.529)
[B]					
	Stock market question $\text{H}\left[P_N(d_{S,1} \geq \underline{d}|Q_{S,1} = 50)\right]$	[0.700,1] (0.567,1)	[0.667,1] (0.517,1)	[0.555,1] (0.420,1)	[0.636,1] (0.555,1)
	Income question $\text{H}\left[P_N(d_{I,1} \geq \underline{d}|Q_{I,1} = 50)\right]$	[0.711,1] (0.536,1)	[0.684,1] (0.489,1)	[0.574,1] (0.388,1)	[0.653,1] (0.546,1)
	Social Security question $\text{H}\left[P_N(d_{B,1} \geq \underline{d}|Q_{B,1} = 50)\right]$	[0.794,1] (0.628,1)	[0.857,1] (0.728,1)	[0.850,1] (0.731,1)	[0.838,1] (0.759,1)
[C]					
	Stock market question $\text{H}\left[P_N(d_{S,1} \geq \underline{d}|Q_{S,1} = \text{M10})\right]$	[0.084,0.551] (0.028,0.651)	[0.141,0.570] (0.061,0.683)	[0.089,0.550] (0.036,0.643)	[0.102,0.556] (0.066,0.614)
	Income question $\text{H}\left[P_N(d_{I,1} \geq \underline{d}|Q_{I,1} = \text{M10})\right]$	[0.061,0.396] (0.007,0.507)	[0.084,0.326] (0.011,0.450)	[0.105,0.487] (0.034,0.604)	[0.083,0.409] (0.045,0.477)
	Social Security question $\text{H}\left[P_N(d_{B,1} \geq \underline{d}|Q_{B,1} = \text{M10})\right]$	[0.145,0.510] (0.052,0.642)	[0.291,0.569] (0.164,0.707)	[0.212,0.560] (0.120,0.672)	[0.213,0.547] (0.153,0.620)

Note: The top entries, i.e., the data in brackets [·], are the estimates of the identification region for the probability of interest. The bottom entries, i.e., the data in parentheses (·), are the estimated asymptotic Bonferroni confidence intervals with level at least 99 percent. For instance, consider the probability that a response to the stock market question is gross rounding in wave 12, i.e., $P(d_{S,1} \geq \underline{d})$. The estimate of the identification region $\text{H}\left[P(d_{S,1} \geq \underline{d})\right]$ is [0.207,0.523], and its estimated Bonferroni 99 percent confidence interval is (0.156,0.586); see Subsection 2.3 and Appendix B for more details.

4   Why do some respondents engage in gross rounding?

This section tries to understand if the probability to provide a gross rounding is related to some covariates. First, one might expect that gross rounding is more likely among respondents with little formal education because the cognitive cost to provide a precise subjective probability or even a small rounding is higher for them. Second, one might expect that younger respondents are more likely to provide gross rounding to the Social Security question. This question elicits the beliefs that respondents have concerning the benefits they will collect when they will turn 70, and younger respondents need to analyze a more complex set of information than older respondents. They are also perhaps less interested in this question than older respondents and put less effort in their answer. Lastly, one may also expect that the probability of providing a gross rounding to the stock market question will be higher in wave 14. Waves 12 and 13 (July-November 1999 and February-May 2000) took place when the return on the Standard & Poor 500 (S&P 500) was above 0.20 per year, while wave 14 (September 2000-March 2001) took place just after a negative shock which was eventually the beginning of a notable bear market.[12]   At that time, respondents might have lacked information to know if the shock was temporary or the beginning of a big decline. Note that it would have been interesting to know if respondents had some financial knowledge or hold assets, but the SEE does not provide such information.

The objective is thus to learn about:

$$\Delta(j, s, t) \equiv P(d_{j,1} \geq \underline{d} | x = s) - P(d_{j,1} \geq \underline{d} | x = t) \tag{7}$$

If $\Delta > 0$ (< 0), a respondent with covariate $x = s$ is $\Delta.100$ percentage points more (less) likely to provide a gross rounding to question j than a respondent with covariate $x = t$. For instance

---

[12]It is usually considered that this bear market began in March 2000. At the beginning of September 2000 (when wave 14 began), the S&P index was over 1500 points (and so was very close to its all-time intraday high of 1552 on 24 March 2000). It was under 1200 points at the end of March 2001 (when wave 14 finished), and bottomed out at 768.63 on October 10, 2002.

$x = s$ may indicate whether a person is less than 40 and $x = t$ if he is over 60 (as mentioned above, this difference is particularly interesting when one considers the Social Security question, i.e., when $j = B$); alternatively $x = s$ may indicate whether a person is interviewed in wave 14 and $x = t$ if he is interviewed in wave 12 (as mentioned above, this difference is particularly interesting when one considers the stock market question, i.e., when $j = S$). The difficulty to learn about Equation 7 is that $\mathrm{P}(d_{j,1} \geq \underline{d}|x = s)$ and $\mathrm{P}(d_{j,1} \geq \underline{d}|x = t)$ are not point-identified under the weak assumptions considered. For a given x, we know that $\mathrm{H}\left[\mathrm{P}(d_{j,1} \geq \underline{d}|x)\right] = [\mathrm{P}(d_{j,1,L} \geq \underline{d}|x), \mathrm{P}(d_{j,1,U} \geq \underline{d}|x)]$. The identification region for the difference between the two probabilities is thus:

$$\mathrm{H}\left[\Delta(j, s, t)\right] = [\, \mathrm{P}(d_{j,1,L} \geq \underline{d}|x = s) - \mathrm{P}(d_{j,1,U} \geq \underline{d}|x = t), \mathrm{P}(d_{j,1,U} \geq \underline{d}|x = s) - \mathrm{P}(d_{j,1,L} \geq \underline{d}|x = t) \,]$$

(8)

To learn about Equation 8, I will estimate $\mathrm{H}\left[\mathrm{P}(d_{j,1} \geq \underline{d}|x = t)\right]$ and $\mathrm{H}\left[\mathrm{P}(d_{j,1} \geq \underline{d}|x = s)\right]$ first. Before to discuss the results, and as a last remark, one may ask if the respondents who do not provide a response have to be included in the sample of interest. One can say that the problem of gross rounding only concerns those who provide a subjective probability. So one can argue that the population of interest is the respondents who provide a subjective probability $Q_{j,1}$. Although put in different terms, the results that one will obtain if he excludes from the analysis the nonresponses are similar to the results that a researcher would obtain if he were assuming that nonresponses are random.[13] As noticed in Subsection 3.1, one can argue however that a nonresponse, in particular a "don't know", is not random: it can be a response strategy for a respondent who has highly imprecise beliefs (because he lacks any relevant information or do not want to put any effort in his answers). If so, a nonresponse is similar to a gross rounding; for a very similar assumption, see also Manski and Molinari (2010, p.223) who consider that a nonresponse is similar to the grossest form of rounding.

I have considered these two cases (i.e., excluding the nonresponses or including them under the

---

[13]Indeed, consider that $d_{j,1}$ is perfectly observed for some respondents, and not at all for others who do not provide any answer. Consider also that nonresponses are random, i.e., $\mathrm{P}(d_{j,1} \geq \underline{d}|Z_j = 1) = \mathrm{P}(d_{j,1} \geq \underline{d}|Z_j = 0)$, where Z_j takes the value one if $d_{j,1}$ is observed, and zero otherwise. It thus implies that the distribution of $d_{j,1}$ coincides with the observable distribution, i.e., $\mathrm{P}(d_{j,1} \geq \underline{d}) = \mathrm{P}(d_{j,1} \geq \underline{d}|Z_j = 1)$.

Table 4: Who is more likely to provide a gross rounding? (based on Assumption 1, with $\alpha = 20$ and $\underline{d} = 30$; Nonresponses excluded)

	Stock market ($j = S$)		Income ($j = I$)		Social Security ($j = B$)	
	$\mathrm{H}_{[\mathrm{P}_N(d_S \geq \underline{d}|x)]}$		$\mathrm{H}_{[\mathrm{P}_N(d_I \geq \underline{d}|x)]}$		$\mathrm{H}_{[\mathrm{P}_N(d_B \geq \underline{d}|x)]}$	
All persons	[0.194,0.533] (0.165,0.569)	(N=1231)	[0.123,0.307] (0.099,0.341)	(N=1247)	[0.294,0.485] (0.254,0.529)	(N=851)
Education						
≤High School	[0.314,0.568] (0.217,0.672)	(N=153)	[0.184,0.343] (0.112,0.431)	(N=195)	[0.374,0.553] (0.261,0.668)	(N=123)
Attended college	[0.216,0.509] (0.162,0.574)	(N=389)	[0.147, 0.332] (0.102,0.391)	(N=413)	[0.341,0.521] (0.265,0.601)	(N=261)
Assoc degree	[0.176,0.519] (0.079,0.647)	(N=102)	[0.097,0.262] (0.022,0.374)	(N=103)	[0.174,0.333] (0.056,0.479)	(N=69)
BA/BS	[0.127,0.526] (0.081,0.595)	(N=355)	[0.083,0.285] (0.043,0.351)	(N=312)	[0.282,0.500] (0.206,0.584)	(N=234)
Master/PhD	[0.088,0.563] (0.025,0.673)	(N=135)	[0.054,0.256] (0.002,0.355)	(N=129)	[0.231,0.451] (0.124,0.577)	(N=104)
Did not know/refused	[0.340,0.567] (0.216,0.696)	(N=97)	[0.147,0.315] (0.053,0.439)	(N=95)	[0.233,0.366] (0.092,0.527)	(N=60)
Age						
18≤Age≤39	[0.197,0.543] (0.152,0.600)	(N=506)	[0.130,0.338] (0.092,0.392)	(N=523)	[0.350,0.623] (0.283,0.691)	(N=337)
40≤Age≤49	[0.152,0.509] (0.096,0.586)	(N=275)	[0.109,0.293] (0.061,0.363)	(N=283)	[0.279,0.474] (0.199,0.562)	(N=211)
50≤Age≤59	[0.169,0.551] (0.103,0.640)	(N=212)	[0.087,0.287] (0.038,0.366)	(N=216)	[0.266,0.388] (0.181,0.482)	(N=180)
Age≥60	[0.260,0.521] (0.187,0.604)	(N=238)	[0.160,0.271] (0.097,0.347)	(N=225)	[0.211,0.268] (0.116,0.371)	(N=123)
Wave						
Wave 12	[0.207,0.523] (0.155,0.586)	(N=415)	[0.127,0.327] (0.084,0.387)	(N=409)	[0.241,0.442] (0.171,0.523)	(N=253)
Wave 13	[0.207,0.517] (0.151,0.587)	(N=342)	[0.120,0.277] (0.076,0.338)	(N=357)	[0.352,0.516] (0.273,0.599)	(N=244)
Wave 14	[0.175,0.553] (0.130,0.611)	(N=474)	[0.122,0.312] (0.084,0.366)	(N=481)	[0.293,0.494] (0.231,0.563)	(N=354)
Age and Education						
Age≥60 and Education≤High School	[0.384,0.589] (0.183,0.793)	(N=39)	[0.244,0.266] (0.079,0.436)	(N=45)	[0.348,0.391] (0.092,0.653)	(N=23)
Age≥60 and Education=Master/PhD	[0.064,0.483] (-0.048,0.715)	(N=31)	[0.031,0.187] (-0.048,0.365)	(N=32)	[0.047,0.095] (-0.071,0.260)	(N=21)
Age≥60 and Education=BA/BS or Master/PhD	[0.160,0.555] (0.055,0.698)	(N=81)	[0.047,0.234] (-0.021,0.371)	(N=64)	[0.093,0.139] (-0.021,0.276)	(N=43)
18≤Age≤39 and Education≤High School	[0.272,0.545] (0.131,0.703)	(N=66)	[0.188,0.388] (0.079,0.524)	(N=85)	[0.367,0.653] (0.190,0.828)	(N=49)
18≤Age≤39 and Education=BA/BS or Master/PhD	[0.114,0.528] (0.055,0.621)	(N=193)	[0.076,0.271] (0.026,0.356)	(N=184)	[0.304,0.601] (0.199,0.713)	(N=128)

Note: The first line in each cell provides the estimated identification region $\mathrm{H}_{[\mathrm{P}_N(d_{j,1} \geq \underline{d}|x)]}$ in brackets, and the number of observations in parentheses (N=.). The bottom entries provide the estimated asymptotic Bonferroni confidence intervals with level at least 99 percent in parentheses.

assumption that they are similar to a gross rounding). In both cases, $\mathrm{H}[\mathrm{P}_N(d_{j,1} \geq \underline{d}|x = s)]$ and

$\mathrm{H}[\mathrm{P}_N(d_{j,1} \geq \underline{d}|x = t)]$ overlap for most variables of interest; hence, the interval given by Equation 8

contains the value zero and it is impossible to make any conclusion. To fully understand, consider

Table 5: Who is more likely to provide gross rounding? (based on Assumption 1, with $\alpha = 20$ and $\underline{d} = 30$; Nonresponses included under the assumption that they are similar to a gross rounding)

	Stock market ($j = S$) $\mathrm{H}_{[\mathrm{P}_N(d_S \geq \underline{d}\|x)]}$		Income ($j = I$) $\mathrm{H}_{[\mathrm{P}_N(d_I \geq \underline{d}\|x)]}$		Social Security ($j = B$) $\mathrm{H}_{[\mathrm{P}_N(d_B \geq \underline{d}\|x)]}$	
All persons	[0.399,0.651] (0.368,0.681)	(N=1651)	[0.337,0.476] (0.308,0.508)	(N=1651)	[0.531,0.657] (0.495,0.692)	(N=1280)
Education						
≤High School	[0.639,0.773] (0.566,0.836)	(N=291)	[0.453,0.560] (0.378,0.635)	(N=291)	[0.624,0.731] (0.537,0.811)	(N=205)
Attended college	[0.414,0.633] (0.359,0.687)	(N=521)	[0.324,0.470] (0.271,0.526)	(N=521)	[0.560,0.680] (0.495,0.741)	(N=391)
Assoc degree	[0.323,0.604] (0.214,0.718)	(N=124)	[0.250,0.387] (0.150,0.499)	(N=124)	[0.430,0.540] (0.302,0.668)	(N=100)
BA/BS	[0.255,0.596] (0.199,0.658)	(N=416)	[0.312,0.463] (0.254,0.527)	(N=416)	[0.508,0.658] (0.439,0.724)	(N=342)
Master/PhD	[0.276,0.653] (0.188,0.747)	(N=170)	[0.282,0.435] (0.193,0.533)	(N=170)	[0.448,0.607] (0.342,0.711)	(N=145)
Did not know/refused	[0.504,0.674] (0.390,0.781)	(N=129)	[0.372,0.496] (0.262,0.609)	(N=129)	[0.525,0.608] (0.395,0.736)	(N=97)
Age						
18≤Age≤39	[0.346,0.628] (0.297,0.678)	(N=621)	[0.267,0.443] (0.221,0.494)	(N=621)	[0.583,0.758] (0.528,0.807)	(N=526)
40≤Age≤49	[0.315,0.603] (0.249,0.671)	(N=340)	[0.258,0.412] (0.197,0.480)	(N=340)	[0.514,0.645] (0.441,0.715)	(N=313)
50≤Age≤59	[0.395,0.673] (0.321,0.744)	(N=291)	[0.323,0.471] (0.252,0.546)	(N=291)	[0.514,0.595] (0.436,0.672)	(N=272)
Age≥60	[0.558,0.714] (0.494,0.772)	(N=399)	[0.526,0.589] (0.461,0.652)	(N=399)	[0.426,0.467] (0.327,0.566)	(N=169)
Wave						
Wave 12	[0.398,0.638] (0.344,0.691)	(N=547)	[0.347, 0.497] (0.294,0.552)	(N=547)	[0.527,0.652] (0.463,0.714)	(N=406)
Wave 13	[0.417,0.645] (0.358,0.702)	(N=465)	[0.324,0.445] (0.268,0.504)	(N=465)	[0.579,0.686] (0.514,0.748)	(N=376)
Wave 14	[0.388,0.668] (0.338,0.716)	(N=639)	[0.339,0.482] (0.291,0.533)	(N=639)	[0.497,0.640] (0.440,0.696)	(N=498)
Age and Education						
Age≥60 and Education≤High School	[0.760,0.840] (0.650,0.934)	(N=100)	[0.660,0.670] (0.538,0.791)	(N=100)	[0.594,0.621] (0.386,0.827)	(N=37)
Age≥60 and Education=Master/PhD	[0.369,0.652] (0.186,0.833)	(N=46)	[0.326,0.434] (0.148,0.623)	(N=46)	[0.231,0.269] (0.017,0.493)	(N=26)
Age≥60 and Education=BA/BS or Master/PhD	[0.414,0.689] (0.296,0.800)	(N=116)	[0.474,0.577] (0.354,0.696)	(N=116)	[0.291,0.327] (0.133,0.490)	(N=55)
18≤Age≤39 and Education≤High School	[0.515,0.696] (0.386,0.816)	(N=99)	[0.303,0.474] (0.184,0.604)	(N=99)	[0.635,0.800] (0.500,0.911)	(N=85)
18≤Age≤39 and Education=BA/BS or Master/PhD	[0.226,0.588] (0.154,0.673)	(N=221)	[0.231,0.393] (0.157,0.478)	(N=221)	[0.541,0.737] (0.449,0.818)	(N=194)

Note: The first line in each cell provides the estimated identification region $\mathrm{H}_{[\mathrm{P}_N(d_{j,1} \geq \underline{d}|x)]}$ in brackets, and the number of observations in parentheses (N=.). The bottom entries provide the estimated asymptotic Bonferroni confidence intervals with level at least 99 percent in parentheses.

Table 4 which presents the results when the nonresponses are excluded. The estimated probability that a respondent with less than a high school diploma provides a gross rounding to the stock market question ($j = S$) is between [0.314,0.568] while the one of a respondent with a BA/BS is between [0.127,0.526]. A respondent with less than a high school diploma is thus between [-0.212,0.441] more

26

likely to provide a gross rounding. In other words, it is impossible to know if respondents with little formal education are more likely to provide a gross rounding.

The sole set of variables which provides clear results are the age variables at the Social Security question: in Table 4, when the nonresponses are excluded, the estimated bound for the probability of providing a gross rounding when a respondent is more than 60 years old is in the interval $[0.211, 0.268]$ (123 respondents); it is in the interval $[0.350, 0.623]$ when a respondent is between 18 and 39 (337 respondents). Thus, application of Equation 8 says that a respondent under 39 is between $[0.082, 0.412]$ more likely to provide a gross rounding. In Table 5, when the nonresponses are included, the estimated bound for the probability of providing a gross rounding when a respondent is more than 60 years old is in the interval $[0.426, 0.467]$ (169 respondents); it is in the interval $[0.583, 0.758]$ when a respondent is between 18 and 39 (337 respondents). Thus, application of 8 says that a respondent under 39 is between $[0.116, 0.332]$ more likely to provide a gross rounding.

Two remarks are in order.

Remark 1. We have also considered some bounds conditioned on more than one covariate. As described in Tables 4 and 5, we have considered the probability that a respondent over 60 with less than a high school diploma provides a gross rounding, the probability that a respondent over 60 with a BA/BS or a Master/PhD provides a gross rounding, and so on.[14] We do find that education matters, in particular at the Social Security question. For instance, when nonrespondents are included, i.e., in Table 5, we do find that the estimated probability that a respondent over 60 with less than a high school diploma provides a gross rounding is between $[0.594, 0.621]$, while it is between $[0.291, 0.327]$ for a respondent over 60 with a BA/BS or a Master/PhD. Thus, respondents over 60 with BA/BS or Master/PhD are between $[0.267, 0.330]$ less likely to provide a gross rounding. Nevertheless, inter-

---

[14]I have merged the categories "Respondents over 60 with a BA/BS" and "Respondents over 60 with a Master or a PhD" because the number of observations in these two categories is small; e.g., only 21 respondents over 60 have a Master or a PhD at the Social Security question in Table 4. I have thus considered the respondents over 60 with a BA/BS or a Master/PhD to mitigate this problem. The number of observations (43) remains however relatively small even when these categories are merged. It illustrates the well-known Curse of Dimensionality inherent to any non-parametric analysis: as we increase the number of attributes, the fewer observations are available to estimate the proportions of interest, so the estimation results may be unreliable (see, e.g., Ahamada and Flachaire, 2010, p.90).

pretation of these estimates should be cautious given the few observations in these categories: among respondents over 60, only 37 respondents have less than a high school diploma and only 55 have a BA/BS or a Master/PhD at the Social Security question.

Remark 2. If the results confirm the idea that younger respondents face more ambiguity than older respondents concerning the Social Security benefits they will collect when they will turn 70 or are perhaps less interested in this question, a reader can be surprised by the few conclusions that we can make. It is important to understand that our inability to conclude is due to the weakness of Assumption 1, and the conservative choices of choosing $\alpha = 20$ and $\underline{d} = 50 - \alpha = 30$. I also considered lower value for α, in particular $\alpha = 10$, and various critical values $\underline{d} = \{10, 15, 20, \ldots, 40\}$ (results are not reported). Beginning the analysis with $\underline{d} = 10$ was not anodyne. The literature usually assumes that the answers other than $(0, 50, 100)$ cannot imply an interval with a width of more than 10 percentage points, and if $\underline{d} = 10$, a response is gross rounding if its degree of imprecision is at least 10 percentage points ($d_{i,j,1} \geq 10$). However, and again, for all $\underline{d} = \{10, 15, 20, \ldots, 40\}$, $\mathrm{H}\left[\mathrm{P}_N(d_{j,1} \geq \underline{d}|x = s)\right]$ and $\mathrm{H}\left[\mathrm{P}_N(d_{j,1} \geq \underline{d}|x = t)\right]$ overlapped for most variables of interest, except, again, for the age variables at the Social Security question.

5 Conclusion

At the end of his Presidential address of the Econometric Society in 1957, Haavelmo (1958, p.357) pointed out:

> "I think most of us feel that if we could use explicitly such variables as, e.g., what peo-
> ple think prices or incomes are going to be [...], we would be able to establish relations
> that could be more accurate and have more explanatory value.[...] It is my belief that if
> we can develop more explicit and a priori convincing economic models in terms of these
> variables, which are realities in the minds of people even if they are not in the current
> statistical yearbooks, then ways and means can and will eventually be found to obtain

actual measurements of such data."

Half a century after Haavelmo's address, it is standard in economic theory to consider that individuals have expectations for unknown outcomes represented by a subjective probability distribution and choose the best alternative in terms of expected utility. An important part of the research in the measurement of expectations has thus elicited expectations in this form, and has taken the responses at face value. The paper clearly shows however that numerous respondents have difficulties to provide probabilities which reflect sharp expressions of beliefs whatever the question. In particular, a majority of 50s and numerous M10 are rounded grossly for all questions.

The method developed in Section 2 and its application in Section 3 cast serious doubts on the way the extent of rounding is usually inferred in particular for the M10. Remark however that the measure of imprecision is applicable only in the context of the SEE two-stage questioning method to elicit subjective probability distributions for continuous outcomes. In the case of a binary discrete variable, there is no suggestive support to elicit in a first stage, so one cannot construct an interval measurement of imprecision. To avoid misclassifications in this case, the only way to infer credibly the extent of rounding would be to ask for ranges of probabilities rather than precise probabilities, as it has been proposed recently by Manski and Molinari (2010, p.229) and Giustinelli and Pavoni (forthcoming). Basically, they ask first the probability that an event will occur, they then ask respondents if they are sure or unsure about their probability. If they are unsure, then they are asked to provide their minimum and maximum chances. Manski and Molinari (2010, p.230) highlight that numerous respondents (248) were willing to report ranges of probabilities, but there is a clear need to validate their approach: they also note that "among the 264 persons who reported that their response was an exact number, almost a quarter (60) reported that their survival probability is precisely 50 percent" (p.330).

In the case of a continuous variable, the elicitation of a sequence of ranges of probabilities would be more complex. Asking for a sequence of ranges can be difficult because monotonicity of a set of (complementary) cumulative distribution functions implies that the lower and upper bounds for a

29

threshold $r_{i,j,2}$, $r_{i,j,3}$, ... or $r_{i,j,K}$ have to be respectively below the lower and upper bounds elicited previously. Given this potential problem with continuous variables, the format of questions in the SEE appears useful: using the measure of imprecision, one can infer more credibly the extent of rounding to the first threshold if it is outside the suggestive support. The same would be true for the last threshold if it is also outside the suggestive support. Concerning the other thresholds, it is important that they belong to the suggestive support. The measures of imprecision at these thresholds will not permit to learn something about the extent of rounding. One can bound however the actual rounding interval at these thresholds using the rounding intervals inferred at the first and the last thresholds: because of monotonicity, the lower and upper bounds of the interval data at the second threshold must be respectively below the lower and upper bounds at the first threshold, while the lower and upper bounds of the interval data at the penultimate threshold must be respectively above the lower and upper bounds at the last threshold; then, by tâtonnement, one can recover a full sequence of ranges. The crucial point is thus to propose an algorithm that insures that the first and last thresholds are outside the suggestive support and that the other ones belong to it.

Lastly, one can ask how one can exploit interval expectations data. The objective of collecting expectations data was to improve the estimation of structural models of choice under uncertainty. The recent papers which estimate models of choice under uncertainty (e.g., Delavande, 2008) combine the data on expectations –that they take at face value– and choices to estimate random utility models in which decision makers are assumed to maximize their subjective expected utility. If a researcher wants to take into consideration the extent of rounding, he needs some econometric tools to deal with interval data, as well as modeling choice behavior. Some tools exist (see, e.g., Manski and Tamer, 2002). But modeling choice behavior when, e.g., rounding is due to a lack of information is perhaps very different than when it is due to the fact that survey respondents lack time and incentive to retrieve a precise probability. If imprecision is due to the fact that survey respondents lack time and incentive when they answer the survey, it is possible that when they make their decision they take

30

time to analyze information and assess their probability precisely. As a result, a researcher might argue that decision makers have a precise subjective probability and make choices which maximize subjective expected utility; however this researcher does not observe precisely each decision maker's precise subjective probability which lies presumably in the interval data inferred preliminary. Consider now that respondents have imprecise probability in mind because they lack information. If a researcher wants to model how decision makers behave when they have an incomplete subjective distribution, the question is then to know which criterion decision makers use. Perhaps they use the maximin rule (i.e., they maximize the minimum expected utility), but other criteria exist (e.g., the more general Hurwicz (1951) criterion). Whatever the answer, there is a clear need to propose a class of models with interval expectations data.

References

Abbas, A. E., Budescu, D. V., Yu, H.-T. and Haggerty, R. (2008), "A comparison of two probability encoding methods: Fixed probability vs. fixed variable values", *Decision Analysis*, vol. 5 nᵒ 4: pp. 190–202.

Ahamada, I. and Flachaire, E. (2010), *Non-Parametric Econometrics*, Oxford University Press.

de Bresser, J. and van Soest, A. (2013), "Survey response in probabilistic questions and its impact on inference", *Journal of Economic Behavior & Organization*, vol. 96: pp. 65–84.

Bruine de Bruin, W., Fischhoff, B. and Halpern-Felsher, B. L. (2000), "Verbal and numerical expressions of probability: "It's a fifty-fifty chance"", *Organizational and Human Decision Processes*, vol. 81 nᵒ 1: pp. 115–131.

Delavande, A. (2008), "Pill, patch, or shot? Subjective expectations and birth control choice", *International Economic Review*, vol. 49 nᵒ 3: pp. 999–1042.

DELAVANDE, A. and KOHLER, H.-P. (forthcoming), "HIV/AIDS-related expectations and risky sexual behavior in Malawi", *Review of Economic Studies.*

DOMINITZ, J. (2001), "Estimation of income expectations models using expectations and realization data", *Journal of Econometrics*, vol. 102 nº 2: pp. 165–195.

DOMINITZ, J. and MANSKI, C. F. (1997), "Using expectations data to study subjective income expectations", *Journal of the American Statistical Association*, vol. 92 nº 439: pp. 855–867.

DOMINITZ, J. and MANSKI, C. F. (2004), "The Survey of Economic Expectations", Technical report, http://faculty.wcas.northwestern.edu/~cfm754/see_introduction.pdf.

DOMINITZ, J. and MANSKI, C. F. (2006), "Measuring pension-benefit expectations probabilistically", *Labour*, vol. 20: pp. 201–236.

GIUSTINELLI, P. and PAVONI, N. (forthcoming), "The evolution of awareness and belief ambiguity in the process of high school track choice", *Review of Economic Dynamics.*

GOURET, F. (2015), "What can we learn from the fifties?", Working paper 20, Thema.

HAAVELMO, T. (1958), "The role of the econometrician in the advancement of economic theory", *Econometrica*, pp. 351–357.

HUDOMIET, P. and WILLIS, R. J. (2013), "Estimating second order probability beliefs from subjective survival data", *Decision Analysis*, vol. 10 nº 2: pp. 152–170.

HURWICZ, L. (1951), "Some specification problems and applications to econometric models", *Econometrica*, vol. 19: pp. 343–344.

KLEINJANS, K. and VAN SOEST, A. (2014), "Rounding, focal point answers and nonresponse to subjective probability questions", *Journal of Applied Econometrics*, vol. 29: pp. 567–585.

LICHTENSTEIN, S., FISCHOFF, B. and PHILLIPS, L. (1982), "Calibration of probabilities: The state of the art to 1980", in KAHNEMAN, D., SLOVIC, P. and TVERSKY, A. (editors), *Judgement Under Uncertainty: Heuristics and Biases*, Cambridge University Press, pp. 306–334.

MANSKI, C. (2003), *Partial Identification of Probability Distributions*, New-York: Springer-Verlag.

MANSKI, C. F. (2004), "Measuring expectations", *Econometrica*, vol. 72 n° 5: pp. 1329–1376.

MANSKI, C. F. and MOLINARI, F. (2010), "Rounding probabilisitic expectations in surveys", *Journal of Business and Economics Statistics*, vol. 28 n° 2: pp. 219–231.

MANSKI, C. F. and TAMER, E. (2002), "Inference of regressions with interval data on a regressor or outcome", *Econometrica*, vol. 70: pp. 519–546.

SAVIN, N. (1984), "Multiple hypothesis testing", in GRILICHES, Z. and INTRILIGATOR, M. D. (editors), *Handbook of Econometrics*, Elsevier, vol. 2, chapter 14, pp. 827–879.

SCHOLLMEYER, G. and AUGUSTIN, T. (2013), "On Sharp Identification Regions for Regression Under Interval Data", Technical Report 143, LMU Institut Für Statistik.

TOURANGEAU, R., RIPS, L. J. and RASINSKI, K. (2000), *The Psychology of Survey Response*, Cambridge University Press.

WALLEY, P. (1991), *Statistical Reasoning with Imprecise Probabilities*, London: Chapman and Hall.

A The Social Security and the income questions

Table 1 in Section 1 describes the stock market question. This appendix describes the two other questions that we will study in Section 3.

A.1 The Social Security benefit question

The format of the Social Security benefit question ($j = B$) described in Table A1 is fundamentally the same as the one of the stock market question described in Table 1: it is the two stage questioning method used in the SEE to elicit subjective probability distributions for continuous outcomes. The first stage asked the respondent to report the lowest and highest possible levels of their future benefits when he will be 70 years old. The responses $r_{i,B,min}$ and $r_{i,B,max}$ to these preliminary questions were then used to set thresholds for four probabilistic questions about the level of benefits, as described in Table A1.

Two remarks are in order. First, Dominitz and Manski (2006, p.213) who analyze these data write that the responses to the preliminary questions were "*used to set thresholds for up to six probabilistic questions about the level of benefits*", and not four as noted in Table A1. To avoid any doubt, this difference is due to the fact that the responses to the preliminary questions were used to set thresholds for four probabilistic questions. But if the response to the fourth probabilistic question $Q_{i,B,4}$ was more than 10 percent, then another question with a higher threshold value would be asked; and if the response to the first probabilistic question $Q_{i,B,1}$ was less than 90 percent, then another question with a lower threshold value would be asked (see Dominitz and Manski, 2004, p.11). This is not crucial for our analysis because we will focus on the response to the first threshold $Q_{i,B,1}$ as explained in the Introduction.

Second, it is important to note that less respondents were interviewed at the Social Security benefit question than at the stock market and income questions. The stock market and the income questions were posed to all the persons interviewed in the three waves of the SEE (persons of age 18 or older).

34

SEE scenario

Suppose you are eligible to collect Social Security benefits when you turn 70. Please think about how much money you would be eligible to collect each year. When considering the dollar value, please ignore the effects of inflation or cost-of-living increases. That is, please respond as if a dollar today is worth the same as a dollar when you turn 70.

Preliminary questions

What do you think is the LOWEST amount of social security benefits, per year, that you would be eligible to receive? (The interviewer entered the response $r_{i,B,min}$ in thousands of Dollars)

What do you think is the HIGHEST amount of social security benefits, per year, that you would be eligible to receive? (The interviewer entered the response $r_{i,B,max}$ in thousands of Dollars)

Algorithm for selection of Social Security benefit thresholds

$\left\lceil \frac{r_{i,B,min}+r_{i,B,max}}{2} \right\rceil$	$r_{i,B,1}$	$r_{i,B,2}$	$r_{i,B,3}$	$r_{i,B,4}$
0 to 19	5	10	15	20
20 to 24	10	15	20	25
25 to 29	15	20	25	30
30 to 34	20	25	30	35
35 to 39	25	30	35	40
40 to 49	30	35	40	50
50 to 59	35	40	50	60
60 to 69	40	50	60	70
70 to 89	50	60	70	80
more than 90	60	80	100	125

Note: The midpoint of the lowest and highest values rounded up to the next integer $\left(\left\lceil \frac{r_{i,B,min}+r_{i,B,max}}{2} \right\rceil \right)$ determined the respondent's four thresholds according to the algorithm presented in this Table.

Sequence of $K = 4$ probabilistic questions

What do you think is the PERCENT CHANCE (or CHANCES OUT OF 100) that you would be eligible to receive over $r_{i,B,k}$ of Social Security benefits per year, when you turn 70? ($Q_{i,B,k}$, $k = 1, 2, 3, 4$, in thousands of Dollars)

But logically the Social Security question concerns only the persons of ages 18-69. Furthermore, before the question described in Table A1, i.e., before the scenario, the preliminary questions and the sequence of probabilistic questions, the SEE began with a description of the Social Security program and a request for the respondent to predict his eligibility for benefits when 70 years old. Only the respondents of ages 18-69 who provided a non-zero probability of eligibility were asked the preliminary questions and the probabilistic questions. For information, the wording of the description of the Social Security program and the eligibility was as follows:

Politicians and the news media have been talking recently about the future of the Social Security retirement system, the federal program providing benefits to retired workers. The amount of benefits for which someone is eligible is currently determined by the person's

retirement age and by earnings prior to retirement. There has been much discussion of changing the form of the Social Security system, so the future shape of the system is not certain. With this in mind, I would like you to think about what kind of Social Security retirement benefits will be available when you are older. In particular, think ahead to when you are about to turn 70 years old and suppose that you are not working at that time. What is the per cent chance that you will be eligible to collect any Social Security retirement benefits at that time?

This question is not important for our analysis because we will focus on the response to the first threshold $Q_{i,B,1}$ as explained in the Introduction. But it is important to understand that the Social Security question described in Table A1 was posed to less respondents than the stock market and income questions. Description of the data and response rates are provided in Subsection 3.1.

A.2   The income question

The format of the income question $(j = I)$ described in Table A1 is almost the same as the one of the stock market question described in Table 1: it is the two stage questioning method used in the SEE to elicit subjective probability distributions for continuous outcomes. The first stage asked the respondent to report the lowest and highest possible levels of their one-year-ahead income. The responses $r_{i,I,min}$ and $r_{i,I,max}$ to these preliminary questions were then used to set thresholds for four probabilistic questions about their income, as described in Table A2.

Two remarks are in order. First, remark in Table A2 that respondents were asked to report the percent chance that their income would be less than a sequence of threshold values. Hence, the probabilistic responses are points on the respondent's subjective cumulative distribution function of their income over the next 12 months. It differs from the stock market and the Social Security benefit questions which asked points on the respondent's subjective complementary cumulative distribution function. So one observes the probabilistic answers $F_{i,I,k} = (100 - Q_{i,I,k}) \equiv [100 - \mathrm{P}\,(R_{i,I} > r_{i,I,k})]$,

$k = 1, 2, \ldots, K$. The method presented in Section 2 considers that the respondents provide a sequence of K points on their subjective complementary cumulative distribution function (i.e., $Q_{i,j,k}$, $k = 1, 2, \ldots, K$). The method also applies for the income question, given that if one observes $F_{i,I,k}$, he can deduce $Q_{i,I,k}$. Hence, it does not change the analysis.

Second, Dominitz (2001, p.177) who analyzes the responses to the income question of Table A2 (but in waves 1 to 4 which took place in the period April 1994-February 1996) writes that the responses to the preliminary questions were used to set thresholds for up to six probabilistic questions about the income, and not four as noted in Table A2. This is also the case in waves 12 to 14 that we analyze in this paper. This difference is due to the fact that the responses to the preliminary questions were used to set thresholds for four probabilistic questions. But if the response to the fourth probabilistic question $F_{i,I,4}$ was less than a 90 percent chance, then another question with a higher threshold value would be asked; and if the response to the first probabilistic question $F_{i,I,1}$ was greater than a 10 percent chance, then another question with a lower threshold value would be asked (see Dominitz and Manski, 2004, p.7). This is not crucial for our analysis because we will focus on the response to the first threshold $F_{i,I,1}$ as explained in the Introduction.

Preliminary questions

What do you think is the LOWEST amount that your OWN total income, from all sources, could possibly be in the next 12 months, BEFORE TAXES? (The interviewer entered the response $r_{i,I,min}$ in thousands of Dollars)

What do you think is the HIGHEST amount that your OWN total income, from all sources, could possibly be in the next 12 months, BEFORE TAXES? (The interviewer entered the response $r_{i,I,max}$ in thousands of Dollars)

Algorithm for selection of income thresholds

$\left\lceil \frac{r_{i,I,min}+r_{i,I,max}}{2} \right\rceil$	$r_{i,I,1}$	$R_{i,I,2}$	$r_{i,I,3}$	$r_{i,I,4}$
0 to 19	10	15	20	25
20 to 24	15	20	25	30
25 to 29	20	25	30	35
30 to 34	25	30	35	40
35 to 39	30	35	40	50
40 to 49	35	40	50	60
50 to 59	40	50	60	70
60 to 69	50	60	70	80
70 to 89	60	70	80	100
more than 90	80	100	125	150

Note: The midpoint of the lowest and highest values rounded up to the next integer $\left(\left\lceil \frac{r_{i,I,min}+r_{i,I,max}}{2} \right\rceil \right)$ determines the respondent's four thresholds according to the algorithm presented in this Table.

Sequence of $K = 4$ probabilistic questions

Still thinking about your OWN total income, BEFORE TAXES, in the next 12 months...
What do you think is the PERCENT CHANCE (or CHANCES OUT OF 100) that your OWN total income, BEFORE TAXES, will be under $r_{i,I,k}$? ($F_{i,I,k} = 100 - Q_{i,I,k}$, $k = 1, 2, 3, 4$, in thousands of Dollars)

B   Asymptotic Bonferroni confidence intervals

Consider that a researcher is interested in the probability θ (e.g., θ can be the probability that a response is gross rounding in the population of interest). $(\widehat{\theta_L}, \widehat{\theta_U})$ are estimates of the bounds (θ_L, θ_U) on the identification region of the partially identified population parameter θ. The objective is to obtain a confidence interval (CI). In order to form the CI, note first that if the sample is large, it follows from the Central Limit Theorem that the standard errors of $\widehat{\theta_L}$ and $\widehat{\theta_U}$ are $S_L = [\theta_L(1-\theta_L)/N]^{0.5}$ and $S_U = [\theta_U(1-\theta_U)/N]^{0.5}$. Consider the one sided CI $(\widehat{\theta_L} - z_{1-(\delta/2)}S_L, +\infty)$ for the lower bound θ_L with coverage $P(\widehat{\theta_L} - z_{1-(\delta/2)}S_L < \theta_L) = 1 - \frac{\delta}{2}$ and the one sided CI $(-\infty, \widehat{\theta_U} + z_{1-(\delta/2)}S_U)$ for the upper bound θ_U with coverage $P(\theta_U < \widehat{\theta_U} + z_{1-(\delta/2)}S_U) = 1 - \frac{\delta}{2}$. The Bonferroni inequality[15] permits to write that $P(\widehat{\theta_L} - z_{1-(\delta/2)}S_L < \theta_L, \theta_U < \widehat{\theta_U} + z_{1-(\delta/2)}S_U) \geq 1-\delta$. It means that $(\widehat{\theta_L} - z_{1-(\delta/2)}S_L, \widehat{\theta_U} + z_{1-(\delta/2)}S_U)$

---

[15] The Bonferroni inequality states that $P(A_1, A_2, \ldots, A_N) \geq 1 - \sum_{n=1}^{N} P(A_n^c)$, where A_n is an event and A_n^c its complement (see, e.g., Savin, 1984, p.834).

has a coverage probability of at least $(1 - \delta)$. This is what is called the Bonferroni asymptotic CI with level at least $(1 - \delta)$ in the main text. Note that if I choose $\delta = 0.01$, so I consider a CI with at least 99 percent, $z_{1-(\delta/2)} \simeq 2.58$. Finally, to obtain the estimated CI, substitute population probabilities by sample estimates to obtain the estimated standard errors $\widehat{S_L} = [\widehat{\theta_L}(1 - \widehat{\theta_L})/N]^{0.5}$ and $\widehat{S_U} = [\widehat{\theta_U}(1 - \widehat{\theta_U})/N]^{0.5}$.

C   Proofs

C.1   Proof of Proposition 1

Equation 6 in Proposition 1 follows Manski (2003, p.18, Proposition 1.4). The relative frequency $P_N(d_{j,1,L} \geq \underline{d}|Q_{j,1} = 50) = \frac{\sum_{i=1}^{N} \mathbf{1}[d_{i,j,1,L} \geq \underline{d}, Q_{i,j,1}=50]}{\sum_{i=1}^{N} \mathbf{1}[Q_{i,j,1}=50]}$ is a possible value of $P_N(d_{j,1} \geq \underline{d}|Q_{j,1} = 50)$ and is dominated by all other possible values of $P_N(d_{j,1} \geq \underline{d}|Q_{j,1} = 50)$. If so, $P_N(d_{j,1,L} \geq \underline{d}|Q_{j,1} = 50)$ is the lower bound on $P_N(d_{j,1} \geq \underline{d}|Q_{j,1} = 50)$. Concerning the relative frequency $P_N(d_{j,1,U} \geq \underline{d}|Q_{j,1} = 50) = \frac{\sum_{i=1}^{N} \mathbf{1}[d_{i,j,1,U} \geq \underline{d}, Q_{i,j,1}=50]}{\sum_{i=1}^{N} \mathbf{1}[Q_{i,j,1}=50]}$, it is also a possible value of $P_N(d_{j,1} \geq \underline{d}|Q_{j,1} = 50)$, but it dominates all other possible values of $P_N(d_{j,1} \geq \underline{d}|Q_{j,1} = 50)$. If so, $P_N(d_{j,1,U} \geq \underline{d}|Q_{j,1} = 50)$ is the upper bound on $P_N(d_{j,1} \geq \underline{d}|Q_{j,1} = 50)$. Given that $P_N(d_{j,1,U} \geq \underline{d}|Q_{i,j,1} = 50) = P_N(d_{j,1,L} \geq \underline{d}|Q_{i,j,1} = 50) + P_N(d_{j,1,U} \geq \underline{d} > d_{j,1,L}|Q_{i,j,1} = 50)$, with $P_N(d_{j,1,U} \geq \underline{d} > d_{j,1,L}|Q_{i,j,1} = 50) = \frac{\sum_{i=1}^{N} \mathbf{1}[d_{i,j,1,U} \geq \underline{d} > d_{i,j,1,L}, Q_{i,j,1}=50]}{\sum_{i=1}^{N} \mathbf{1}[Q_{i,j,1}=50]}$, we obtain Equation 6.

Now, if $Q_{i,j,1} = 50$ and $r_{i,j,1} \notin (r_{i,j,min}, r_{i,j,max})$, then $d_{i,j,1,L} = 50 - \alpha$ and $d_{i,j,1,U} = 50$ according to Equations 2 and 3, so $d_{i,j,1,L} = 50 - \alpha \geq \underline{d}$ for all $\underline{d} \leq 50 - \alpha$. If $Q_{i,j,1} = 50$ and $r_{i,j,1} \in (r_{i,j,min}, r_{i,j,max})$, then $d_{i,j,1,L} = 0$ and $d_{i,j,1,U} = 50$ according to Equation 4, so $d_{i,j,1,U} = 50 > \underline{d} > d_{i,j,1,L} = 0$ for all $\underline{d} \leq 50 - \alpha$. If so, $P_N(d_{j,1,L} \geq \underline{d}|Q_{j,1} = 50) = \frac{\sum_{i=1}^{N} \mathbf{1}[d_{i,j,1,L} \geq \underline{d}, Q_{i,j,1}=50]}{\sum_{i=1}^{N} \mathbf{1}[Q_{i,j,1}=50]} = \frac{\sum_{i=1}^{N} \mathbf{1}[r_{i,j,1} \notin (r_{i,j,min}, r_{i,j,max}), Q_{i,j,1}=50]}{\sum_{i=1}^{N} \mathbf{1}[Q_{i,j,1}=50]}$ for all $\underline{d} \leq 50 - \alpha$; and $P_N(d_{j,1,U} \geq \underline{d} > d_{j,1,L}|Q_{i,j,1} = 50) = \frac{\sum_{i=1}^{N} \mathbf{1}[d_{i,j,1,U} \geq \underline{d} > d_{i,j,1,L}, Q_{i,j,1}=50]}{\sum_{i=1}^{N} \mathbf{1}[Q_{i,j,1}=50]} = \frac{\sum_{i=1}^{N} \mathbf{1}[r_{i,j,1} \in (r_{i,j,min}, r_{i,j,max}), Q_{i,j,1}=50]}{\sum_{i=1}^{N} \mathbf{1}[Q_{i,j,1}=50]}$ for all $\underline{d} \leq 50 - \alpha$.∎

C.2 Proof of Proposition 2

To show that $P_N(d_{j,1,L} \geq \underline{d}) = \frac{\sum_{i=1}^{N} \mathbf{1}[d_{i,j,1,L} \geq \underline{d}]}{N}$ is identical for all combinations of $\underline{d}$ and α which satisfy $\underline{d} = 50 - \alpha$, consider the three following cases: (i.) $r_{i,j,1} \leq r_{i,j,min}$, (ii.) $r_{i,j,1} \geq r_{i,j,max}$, and (iii.) $r_{i,j,1} \in (r_{i,j,min}, r_{i,j,max})$.

(i.) If $r_{i,j,1} \leq r_{i,j,min}$, then $d_{i,j,1,L} = \max\{0, 100 - Q_{i,j,1} - \alpha\}$ (see Equation 2). If $\underline{d} = 50 - \alpha$, then $d_{i,j,1,L} = \max\{0, 100 - Q_{i,j,1} - \alpha\} \geq 50 - \alpha$ when $Q_{i,j,1} \leq 50$. So $d_{i,j,1,L} \geq \underline{d}$ if $Q_{i,j,1} \leq 50$ for all combinations of $\underline{d}$ and α which satisfy $\underline{d} = 50 - \alpha$.

(ii.) If $r_{i,j,1} \geq r_{i,j,max}$, $d_{i,j,1,L} = \max\{0, Q_{i,j,1} - \alpha\}$ (see Equation 3). If $\underline{d} = 50 - \alpha$, then $d_{i,j,1,L} = \max\{0, Q_{i,j,1} - \alpha\} \geq 50 - \alpha$ when $Q_{i,j,1} \geq 50$. So $d_{i,j,1,L} \geq \underline{d}$ if $Q_{i,j,1} \geq 50$ for all combinations of $\underline{d}$ and α which satisfy $\underline{d} = 50 - \alpha$.

(iii.) Lastly, when $r_{i,j,1} \in (r_{i,j,min}, r_{i,j,max})$, $d_{i,j,1,L} = 0$ and $d_{i,j,1,U} \geq 50$ (see Equation 4). Hence, when $r_{i,j,1} \in (r_{i,j,min}, r_{i,j,max})$, $d_{i,j,1,L} = 0 < \underline{d}$ for all combinations of $\underline{d}$ and α which satisfy $\underline{d} = 50 - \alpha$.

We thus obtain that

$$P_N(d_{j,1,L} \geq \underline{d}) = \frac{\sum_{i=1}^{N} \mathbf{1}\left[r_{i,j,1} \leq r_{i,j,min}, Q_{i,j,1} \leq 50\right]}{N} + \frac{\sum_{i=1}^{N} \mathbf{1}\left[r_{i,j,1} \geq r_{i,j,max}, Q_{i,j,1} \geq 50\right]}{N}$$

for all combinations of $\underline{d}$ and α which satisfy $\underline{d} = 50 - \alpha$.∎

D   Nonresponses

This appendix studies if nonresponses are related to some attributes. Table D1 shows that the relative frequency of nonresponses is higher among less educated persons: in particular at the stock market question, it is falling from 47.4 percent for those with high school diploma or less to 14.6 percent for those with a BS/BA; this difference is statistically significant at the 0.01 level according to a z-test ($z = 9.63$). Nonresponses are also related to age but in very different ways according to the question. At the stock market and the income questions, older respondents are more likely not

to respond. For instance, at the stock market question, 40.3 percent of the respondents over 60 did not respond versus 18.5 percent of the respondents under 39; this difference is statistically significant at the 0.01 level ($z = 7.49$). On the contrary at the Social Security question, older respondents are more likely to respond. 27 percent of the respondents over 60 did not respond versus 36 percent of those under 39; this difference is statistically significant at a level of 0.05, but not at the 0.01 level ($z = 2.17$). To describe how the probability of a nonresponse varies with multiple personal attributes, I also estimated probit models. The univariate patterns evident in Table D1 recurred.

Table D1: Frequency of nonresponses by attribute

		Stock market ($j = S$)		Income ($j = I$)		Social Security ($j = B$)	
All persons		0.254 (0.011)	(N=1651)	0.244 (0.010)	(N=1651)	0.335 (0.013)	(N=1280)
Education							
	≤High school	0.474 (0.029)	(N=291)	0.329 (0.027)	(N=291)	0.4 (0.034)	(N=205)
	Attended college	0.253 (0.019)	(N=521)	0.207 (0.017)	(N=521)	0.332 (0.024)	(N=391)
	Assoc degree	0.177 (0.034)	(N=124)	0.169 (0.033)	(N=124)	0.310 (0.046)	(N=100)
	BA/BS	0.146 (0.017)	(N=416)	0.250 (0.021)	(N=416)	0.315 (0.025)	(N=342)
	Master/PhD	0.206 (0.031)	(N=170)	0.241 (0.033)	(N=170)	0.282 (0.037)	(N=145)
	Did not know/refused	0.248 (0.038)	(N=129)	0.263 (0.038)	(N=129)	0.381 (0.049)	(N=97)
Age							
	18≤Age≤39	0.185 (0.015)	(N=621)	0.157 (0.014)	(N=621)	0.359 (0.021)	(N=526)
	40≤Age≤49	0.191 (0.021)	(N=340)	0.167 (0.020)	(N=340)	0.326 (0.026)	(N=313)
	50≤Age≤59	0.271 (0.026)	(N=291)	0.257 (0.025)	(N=291)	0.338 (0.028)	(N=272)
	Age≥60	0.403 (0.024)	(N=399)	0.436 (0.025)	(N=399)	0.272 (0.034)	(N=169)
Wave							
	Wave 12	0.241 (0.018)	(N=547)	0.252 (0.018)	(N=547)	0.377 (0.024)	(N=406)
	Wave 13	0.264 (0.020)	(N=465)	0.232 (0.019)	(N=465)	0.351 (0.024)	(N=376)
	Wave 14	0.258 (0.017)	(N=639)	0.247 (0.017)	(N=639)	0.289 (0.020)	(N=498)

Notes: i. The first line in each cell provides the probability estimate of a nonresponse conditional on the covariate considered, and the number of observations in parentheses (N=.). As an example, consider the stock market question. The first line ("All persons") says that of the 1651 persons who were interviewed, 25.4 percent did not provide a response. ii. The bottom entries provide the asymptotic standard errors of the sample estimates in parentheses.