ECONSTOR
Make Your Publications Visible.

A Service of

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Kleinjans, Kristin J.; van Soest, Arthur

Working Paper

Nonresponse and focal point answers to subjective probability questions

IZA Discussion Papers, No. 5272

Provided in Cooperation with:
IZA – Institute of Labor Economics

Suggested Citation: Kleinjans, Kristin J.; van Soest, Arthur (2010) : Nonresponse and focal point answers to subjective probability questions, IZA Discussion Papers, No. 5272, Institute for the Study of Labor (IZA), Bonn

This Version is available at:
https://hdl.handle.net/10419/46122

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

DISCUSSION PAPER SERIES

IZA DP No. 5272

Nonresponse and Focal Point Answers to Subjective Probability Questions

Kristin J. Kleinjans
Arthur van Soest

October 2010

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Nonresponse and Focal Point Answers to Subjective Probability Questions

Kristin J. Kleinjans
California State University, Fullerton

Arthur van Soest
Netspar, Tilburg University,
RAND and IZA

Discussion Paper No. 5272
October 2010

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions.

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

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

IZA Discussion Paper No. 5272
October 2010

ABSTRACT

Nonresponse and Focal Point Answers to Subjective Probability Questions*

We develop and estimate a panel data model explaining the answers to questions about subjective probabilities, using data from the US Health and Retirement Study. We explicitly account for nonresponse, rounding, and focal point "50 percent" answers. Our results indicate that for three of the four questions considered, almost all 50 percent answers can be explained by rounding. We also find observed and unobserved heterogeneity in the tendencies to report rounded values or a focal answer, explaining persistency in 50 percent-answers over time. Incorporating rounding and focal answers changes some of the conclusions about the socio-economic factors that determine expectations.

JEL Classification: C81, D84

Keywords: item nonresponse, rounding, expectations

Corresponding author:

Kristin J. Kleinjans
Department of Economics
Mihaylo College of Business and Economics
California State University – Fullerton
Fullerton, CA 92834-6848
USA
E-mail: kkleinjans@fullerton.edu

* Kleinjans thanks the Research Fund of Aarhus University for financial support. We would like to thank Dimitris Georgarakos, Michael Hurd, Pierre-Carl Michaud, Daniel Schunk, Federica Teppa, and participants at the CeRP Workshop in Turin and the Netspar Pension Workshop in Amsterdam for helpful comments on an earlier draft of the paper.

1 Introduction

Expectations play a crucial role in many economic models. The recent literature on measuring expectations of individuals proposes to use subjective probability questions (see, e.g., Manski 2004). Answers to such questions are used more and more frequently to understand if expectations and outcomes are related in a systematic manner, to evaluate if individual behavior changes in response to changing expectations, and to relax assumptions on expectation formation in models with forward looking agents.[3] This results in an increased need for a better understanding of response behavior to such questions, especially of item nonresponse and potential focal point answers.

In this paper, we study the response patterns to four questions about subjective probabilities in the Health and Retirement Study (HRS). Specifically, we are interested in disentangling to which extent the observed response patterns are driven by the genuine underlying probability, by rounding, and by a tendency to give focal point answers. Moreover, we want to analyze how these tendencies vary with observed and unobserved characteristics and how accounting for rounding and focal point answers affects the estimates of the determinants of expectations.

We develop and estimate a panel data model where the response to a question about subjective probabilities is a two-step process. In the first step, an individual chooses either not to respond to a question at all, to give a focal point answer unrelated to the probability of interest, or to give an answer that is the exact or rounded value of the true probability. In the latter case, the second step determines the degree of rounding and thus the reported answer. Answers of "50 percent" (50/50 answers) can be purely focal point answers (that is, essentially equivalent to non-response), as well as the result of rounding.

Item nonresponse to a question can be the result of several underlying causes.[4] It can be an expression of the fact that the answer is not known, that the concept of the answer is not

[3] See Bernheim (1990) and Manski (2004) for overviews of subjective expectations and their usage in economic research, and Dominitz and Manski (2005) for an overview of surveys eliciting subjective expectations. For an example of research using expectations data to relax assumptions on expectation formation, see Kapteyn et al. (2009).

[4] Our sample is too small to distinguish answers of "don't know" and refusals. See Shoemaker et al. (2002) for the differences in determinants of these two types of nonresponse.

understood, of uncertainty about the item in question, or a refusal to provide an answer. All of these circumstances are likely to be non-random. A focal point answer is a possible alternative to item nonresponse, and may be seen as more socially desirable than a "don't know" or "refuse". There is evidence that answers of "50 percent" in response to questions about probabilities reflect high uncertainty and the meaning of a "50/50 chance" in daily language and do not necessarily indicate a 50% probability (see, for example, Bruine de Bruin et al., 2000). Such answers might also reflect ambiguity in the sense that the probability of the potential outcome cannot be readily assessed on the basis of available evidence (Fox and Tversky, 1995). Hence, in the first step of our framework, an individual decides whether to give a probability, to say "don't know" or "refuse", or to give a "50" answer in the sense described above.

In the second step, an individual who had chosen to give a numerical answer decides on the number to give. Commonly, answers to subjective probability questions are heaped to different degrees at multiples of 5 (see, e.g., Manski and Molinari, 2010). We interpret this as rounding. Rounding could be a consequence of uncertainty, but could also be related to imprecision in reporting or thinking about probabilities (Kézdi and Willis, 2009). In the empirical model, we take into account that different individuals may tend to use different degrees of rounding.

The literature studying focal point answers and patterns in responses to subjective probability questions is small. Lillard and Willis (2001) and Kézdi and Willis (2009) study focal point answers and use them to construct indicators of precision of probabilistic thinking. They find that stock ownership and the fraction of risky assets in a wealth portfolio are positively related to such an index. Bassett and Lumsdaine (2001) illustrate the existence of systematic errors in subjective responses to probability questions by identifying a common component across subjective responses that is unrelated to the specific question. Their findings suggest that benchmarking subjective probabilities given by individuals who might have had problems understanding the probability question improves inference. We expand this line of research by developing a model that can account for and disentangle item nonresponse, focal point answers, and rounding.

This paper also differs from existing studies by exploiting the panel nature of the data: the same respondents were asked the same subjective probability questions seven times over a period of 12 years. Using random effects models, we allow for unobserved factors that make some respondents more likely not to respond at all, to give a 50/50 focal answer, or to round

excessively, e.g. to multiples of 25 or 50. The various random effects are allowed to be correlated so that, for example, the model can capture that respondents who tend not to respond may also have a large tendency to round.

We use data from the Health and Retirement Study (HRS), which includes a variety of probability questions, making it suitable for this study. In addition, its large sample size, the number of waves available (the HRS was started in 1992 and is repeated every two years), and its national representativeness for older age groups (ages 50 and older) allow in depth study. We use four questions about subjective probabilities, which are virtually unchanged over the years: the probabilities of receiving an inheritance, of leaving a bequest of USD 10,000 or more, of working full-time past age 62, and of having a work-limiting health condition within the next 10 years. The questions on inheritances and bequests have been used in investigations of the effects of wealth shocks on retirement (Brown, Coile and Weisbenner, 2010), on bequests (Hurd and Smith, 2002), optimal savings (Scholz, Seshadri and Khitatrakun, 2006), bequest motives and precautionary savings (Mok, 2010) and long-term insurance purchasing behavior (Cramer and Jensen, 2006). The work related questions have been used to analyze the effects of old age social security reforms (Michaud and van Soest, 2008), reverse retirement (Maestas, 2010) and to forecast labor force participation (Michaud and Rohwedder, 2008).

Our paper is related to Manski and Molinari (2010) who focus on rounding and, using the concept of identification up to a bounding interval, show that rounding can seriously limit the information on the means of the true underlying probabilities given covariates. With minimal assumptions on the rounding process, they find large and uninformative bounding intervals for the conditional expectations of interest for subjective survival probabilities in the HRS. Our analysis is different in many respects. First, we not only consider rounding, but also focal point answers and selective item nonresponse. Second, we look at one probability at the time but exploit the panel nature of the data with repeated observations of the same probability for the same respondents. This allows us to use parametric panel data models that point-identify the parameters of interest, including rounding probabilities and focal point probabilities. We also explain intuitively how our model identifies the rounding probabilities and the probability of a focal point 50/50 answer from the peaked nature of the distribution of the reported probabilities.

For three of the four expectation variables under study, we find that most 50 percent answers are related to rounding and not to focal point answers. The exception is the question

on the probability of future work related health problems, where the number of 50 percent-answers is so large compared to the number of 0 percent- and 100 percent-answers that rounding cannot explain it. Rounding and nonresponse are strongly related to socio--demographic variables, including race and ethnicity. When comparing the results from our model with those of a random-effects tobit model not accounting for rounding or focal points, we find that most but not all coefficients are similar. Notwithstanding this similarity, not taking into account the possibility of rounding, focal points answers and item nonresponse leads to an overestimation of the effects of socio-demographic variables on the expected probabilities to leave a bequest or receive an inheritance.

In the remainder of the paper, we start with an overview of possible psychological reasons for item nonresponse, focal point answers and rounding, closely following Schwarz and Oyserman (2001) (Section 2). We then describe the empirical model (Section 3). After a description of the data and variables used, including descriptive statistics of the frequency and patterns of item nonresponse and 50/50 answers (Section 4), we present the estimation results in Section 5. Section 6 concludes.

2 Reasons for Item Nonresponse, Focal Point Answers and Rounding

It is useful to think about the process of responding to a question as consisting of different steps. The steps are, however, not completely separable and should only be thought of as a tool to better understand the processes involved in responding to a survey question. Schwarz and Oysermann (2001) distinguish five different steps in the respondent's answering process: 1. Understanding, 2. Recalling, 3. Inferring, 4. Mapping, and 5. Editing. All of these steps influence the likelihood of item nonresponse, focal point answers and rounding.[5]

In the first step, the respondent has to understand and interpret the meaning of the question. This step is influenced by the context of the question and the survey in general, but also by the respondent's knowledge of and interest in the topic. Questions about probabilities require comprehension of the concept of probability and an interpretation of the meaning of the question. For example, a question about the probability of working full-time past age 62 requires an interpretation of what is meant by "full-time" and by "work", which, e.g., may

[5] See Schwarz and Oyserman (2001) for details. For a different characterization of the response process resulting in similar determinants of item nonresponse, see Beatty and Hermann (2002). See also Tourangeau et al. (2000) for a more detailed discussion of survey response.

include unpaid work or not. The relevance of the question also varies across respondents. For a 60 years old respondent, the question about working at age 62 will have greater relevance than for a younger respondent. Respondents who have difficulty understanding the question or for whom the topic is less relevant may be more likely to answer "don't know" or "refuse", but may also be more uncertain and round more, or be more likely to give a focal 50/50 answer.

The second step requires the respondent to recall relevant behavior and information. In the earlier example of the probability of working full-time past age 62, retrieving relevant information would include recalling the current employer's retirement age policies. Respondents who have trouble recalling the relevant information will be less likely to give a (non-focal) numerical answer and will more often answer "50 percent" and thereby express uncertainty.

In the third step, the respondent has to make inferences about the answer based upon her understanding of the question and recalling of relevant behavior and information. This is likely to be influenced by the response alternatives, including the possibility to express uncertainty, especially if frequency scales are involved. It can also be influenced by previous questions (order effects), for example because the respondent recalled related issues. Again, if inference is difficult, nonresponse, a 50/50 focal answer, or rounding may be more likely.

In the fourth step, respondents have to map their answer onto the response format, that is, the response alternatives given. If the respondent lacks knowledge and there is no possibility to express uncertainty, the respondent might choose not to respond at all. One can also think about this as a matter of precision; if the possible outcomes are too far apart and none of the events is much more likely than the other, respondents may not feel certain enough to choose a specific numerical answer. The less adequate the response format for the answer the respondent has in mind, the higher the likelihood for item nonresponse and (probably) also for a focal 50/50 answer.

The fifth and last step of responding to a question is the editing stage. At this point, the respondent decides which information to give. The more sensitive the subject to the respondent, the less likely she is to give an answer. This editing also depends on social norms and the willingness to admit ignorance. When choosing not to respond, individuals may rather answer "50/50" rather than "don't know" or "refuse" since this may be more socially acceptable or seen as evoking less conflict.

3 A Model of Item Nonresponse, Focal Point Answers and Rounding

In this section we introduce a panel data model explaining the reported subjective probabilities, explicitly accounting for rounding, focal point answers and item non-response. The model combines ideas of the rounding model of Heitjan and Rubin (1990) with those of models that account for misclassified discrete dependent variables (see, for example, Hausman et al. (1998) for cross-section data and Dustmann and van Soest (2001) for panel data). We model one subjective probability at a time and will apply the model separately to the four subjective probabilities discussed in Section 1.

The true (unobserved) subjective probability is assumed to be driven by a latent variable $y_{i,t}^*$, modeled as follows:

$$y_{i,t}^* = x_{i,t}\beta + \alpha_i + \varepsilon_{i,t} .$$

Here $\varepsilon_{i,t}$ is an error term and α_i is a respondent-specific unobserved heterogeneity term that does not vary over time, reflecting, e.g., unobserved factors that make the event referred to in the question less or more likely, or the degree of optimism or pessimism of respondent *i*. The vector $x_{i,t}$ is a set of (strictly exogenous) explanatory variables observed in wave *t* and β is a vector of unknown parameters to be estimated. Distributional assumptions on $\varepsilon_{i,t}$ and α_i will be given below. The (usually unobserved) true probability $y_{i,t}^T$ is the value of $y_{i,t}^*$, censored at lower bound 0 and upper bound 100:

$$y_{i,t}^T = \max(0, \min(y_{i,t}^*, 100)) .$$

The subjective probability (in %) reported by respondent *i* in wave *t* is denoted by $y_{i,t}$, where $y_{i,t}$ is missing if respondent *i* answers "don't know" or "refuse" and is an integer number between 0 and 100 otherwise. Irrespective of censoring and focal points, it seems obvious from the peaked nature of the distribution of the raw data (see the histograms in Section 4), that the observed probabilities are characterized by rounding. Observed probabilities (in %) are usually a multiple of 10 or 25, sometimes another multiple of 5, and only occasionally not a multiple of 5. In order to take this into account we will explicitly model the rounding process, allowing for rounding to multiples of 50, 25, 10, 5 or 1. Rounding to a multiple of 50 is the maximum form of rounding we consider, leading to answers of 0, 50 or 100; rounding to a multiple of 1 is the minimum extent of rounding allowed for – the survey design did not allow for other answers than integers. Rounding to

multiples of 25, 10 or 5 can be seen as intermediate cases. The (partially observed) type of rounding is denoted by $R_{i,t}$:

$R_{i,t} = 1$: The probability is rounded to a multiple of 1

$R_{i,t} = 2$: The probability is rounded to a multiple of 5

$R_{i,t} = 3$: The probability is rounded to a multiple of 10

$R_{i,t} = 4$: The probability is rounded to a multiple of 25

$R_{i,t} = 5$: The probability is rounded to a multiple of 50.

Because of the plausible way in which these forms of rounding can be ordered, we will use an ordered response equation to explain which form of rounding is used. This equation is given by:

$$R_{i,t}^{*} = x_{i,t}\beta^{R} + \alpha_{i}^{R} + \varepsilon_{i,t}^{R}$$

$$R_{i,t} = j \text{ if } m_{j-1} < R_{i,t}^{*} \leq m_{j},\ j = 1,...,5 .$$

Here $\varepsilon_{i,t}^{R}$ is an error term and α_{i}^{R} is a respondent-specific unobserved heterogeneity term that does not vary over time, reflecting unobserved respondent characteristics that drive the respondent's extent of rounding; β^{R} is a vector of unknown parameters and $m_1,...,m_4$ are unknown (auxiliary) parameters also. By definition, $m_0 = -\infty$ and $m_5 = \infty$.

Instead of reporting a rounded value of the true subjective probability, respondents may also decide not to report any value ("don't know" / "refuse"), or may decide to simply respond 50 as an expression of complete uncertainty. In this latter case we observe $y_{i,t} = 50$. We call this a "50/50" answer. It does not stem from rounding – even in a case where the true probability is closer to 0 or to 100 than to 50, such a 50/50 answer can be given.

We will model the type of answer $D_{i,t}$ using a random effects multinomial logit model with three possible outcomes:

$$u_{i,t}^{j} = x_{i,t}\beta^{j} + \alpha_{i}^{j} + \varepsilon_{i,t}^{j},\ j = 1,2,3$$

$$D_{i,t} = j \text{ if } u_{i,t}^{j} \geq u_{i,t}^{k},\ k = 1,2,3$$

$\beta^1 = 0;\ \alpha_i^1 = 0$

$\varepsilon_{i,t}^j$ *i.i.d.* GEV(I) independent of α_i^2, α_i^3 and $x_{i,t}$.

The benchmark outcome is $D_{i,t} = 1$: rounding; the other outcomes are "don't know" or "refuse" ($D_{i,t} = 2$) and "50/50" ($D_{i,t} = 3$). The assumptions $\beta^1 = 0; \alpha_i^1 = 0$ are normalizations without loss of generality. The distributional assumptions[6] on the error terms $\varepsilon_{i,t}^j$ lead to the following multinomial logit probabilities, conditional on observed characteristics $x_{i,t}$ and unobserved characteristics α_i^2 and α_i^3:

$$P(D_{i,t} = j \mid x_{i,t}, \alpha_i^2, \alpha_i^3) = \exp(x_{i,t}\beta^j + \alpha_i^j) / \sum_{k=1}^{3} \exp(x_{i,t}\beta^k + \alpha_i^k);\ j = 1,2,3.$$

In addition to the assumptions on $\varepsilon_{i,t}^j$, $j = 1,...,3$, given above, the distributional assumptions on error terms and (random) individual effects are as follows:

$\varepsilon_{i,t}$ *i.i.d.* $N(0, \sigma_\varepsilon^2)$, independent of $\varepsilon_{i,s}^1, ..., \varepsilon_{i,s}^3$ and $x_{i,s}, s = 1,...,T$.

$\varepsilon_{i,t}^R$ *i.i.d.* N(0,1), independent of $\varepsilon_{i,s}, \varepsilon_{i,s}^1, ..., \varepsilon_{i,s}^3$ and $x_{i,s}, s = 1,...,T$.

$(\alpha_i, \alpha_i^R, \alpha_i^2, \alpha_i^3)'$ *i.i.d.* $N(0, \Sigma_\alpha)$, independent of $x_{i,s}, \varepsilon_{i,s}$, $\varepsilon_{i,s}^R$, and $\varepsilon_{i,s}^1, ..., \varepsilon_{i,s}^3$, $s = 1,...,T$.

Here σ_ε^2 is a parameter to be estimated, and Σ_α is written as $\Sigma_\alpha = \Lambda\Lambda'$, where Λ is a positive semi-definite lower diagonal matrix with parameters to be estimated, so that any arbitrary covariance matrix of the individual effects Σ_α is allowed for.

The model is estimated by maximum simulated likelihood. For a given respondent *i* the conditional likelihood contribution $L_i^c = L_i^c(\alpha_i, \alpha_i^R, \alpha_i^2, \alpha_i^3)$ given $(\alpha_i, \alpha_i^R, \alpha_i^2, \alpha_i^3)$ can be computed straightforwardly. We present it in detail since this helps to understand the nature of

[6] GEV(I) is the Type I Generalized Extreme Value distribution, also known as the Weibüll distribution: $P(\varepsilon_{i,t}^j \le z) = \exp(-\exp(-z))$.

the model; it is essentially a panel data version of a model with endogenous regime switching and unobserved regimes (see, e.g., Quandt and Ramsey, 1978):[7]

$$L_i^c(\alpha_i,\alpha_i^R,\alpha_i^2,\alpha_i^3)=\prod_{t=1}^{T}L_{i,t}^c(\alpha_i,\alpha_i^R,\alpha_i^2,\alpha_i^3)$$

with:

$$L_{i,t}^c(\alpha_i,\alpha_i^R,\alpha_i^2,\alpha_i^3)=P(D_{i,t}=2\mid x_{i,t},\alpha_i^2,\alpha_i^3) \text{ if } y_{i,t} \text{ is not reported}$$

$$L_{i,t}^c(\alpha_i,\alpha_i^R,\alpha_i^2,\alpha_i^3)=P(D_{i,t}=1\mid x_{i,t},\alpha_i^2,\alpha_i^3)P(R_{i,t}=1\mid x_{i,t},\alpha_i^R)P(y_{i,t}-0.5<y_{i,t}^*\le y_{i,t}+0.5\mid x_{i,t},\alpha_i)$$
$$\text{if } y_{i,t}\in\{0,1,...,100\} \text{ and } y_{i,t}\notin\{5,10,15,....,100\}$$

$$L_{i,t}^c(\alpha_i,\alpha_i^R,\alpha_i^2,\alpha_i^3)=P(D_{i,t}=1\mid x_{i,t},\alpha_i^2,\alpha_i^3)[P(R_{i,t}=1\mid x_{i,t},\alpha_i^R)P(y_{i,t}-0.5<y_{i,t}^*\le y_{i,t}+0.5\mid x_{i,t},\alpha_i)+$$
$$+P(R_{i,t}=2\mid x_{i,t},\alpha_i^R)P(y_{i,t}-2.5<y_{i,t}^*\le y_{i,t}+2.5\mid x_{i,t},\alpha_i)]$$
$$\text{if } y_{i,t}\in\{5,15,35,45,55,65,85,95\}$$

$$L_{i,t}^c(\alpha_i,\alpha_i^R,\alpha_i^2,\alpha_i^3)=P(D_{i,t}=1\mid x_{i,t},\alpha_i^2,\alpha_i^3)[P(R_{i,t}=1\mid x_{i,t},\alpha_i^R)P(y_{i,t}-0.5<y_{i,t}^*\le y_{i,t}+0.5\mid x_{i,t},\alpha_i)+$$
$$+P(R_{i,t}=2\mid x_{i,t},\alpha_i^R)P(y_{i,t}-2.5<y_{i,t}^*\le y_{i,t}+2.5\mid x_{i,t},\alpha_i)+$$
$$+P(R_{i,t}=3\mid x_{i,t},\alpha_i^R)P(y_{i,t}-5<y_{i,t}^*\le y_{i,t}+5\mid x_{i,t},\alpha_i)]$$
$$\text{if } y_{i,t}\in\{10,20,30,40,60,70,80,90\}$$

$$L_{i,t}^c(\alpha_i,\alpha_i^R,\alpha_i^2,\alpha_i^3)=P(D_{i,t}=1\mid x_{i,t},\alpha_i^2,\alpha_i^3)[P(R_{i,t}=1\mid x_{i,t},\alpha_i^R)P(y_{i,t}-0.5<y_{i,t}^*\le y_{i,t}+0.5\mid x_{i,t},\alpha_i)+$$
$$+P(R_{i,t}=2\mid x_{i,t},\alpha_i^R)P(y_{i,t}-2.5<y_{i,t}^*\le y_{i,t}+2.5\mid x_{i,t},\alpha_i)+$$
$$+P(R_{i,t}=4\mid x_{i,t},\alpha_i^R)P(y_{i,t}-12.5<y_{i,t}^*\le y_{i,t}+12.5\mid x_{i,t},\alpha_i)]$$
$$\text{if } y_{i,t}\in\{25,75\}$$

$$L_{i,t}^c(\alpha_i,\alpha_i^R,\alpha_i^2,\alpha_i^3)=P(D_{i,t}=1\mid x_{i,t},\alpha_i^2,\alpha_i^3)[P(R_{i,t}=1\mid x_{i,t},\alpha_i^R)P(y_{i,t}^*\le 0.5\mid x_{i,t},\alpha_i)+$$
$$+P(R_{i,t}=2\mid x_{i,t},\alpha_i^R)P(y_{i,t}^*\le 2.5\mid x_{i,t},\alpha_i)+$$
$$+P(R_{i,t}=3\mid x_{i,t},\alpha_i^R)P(y_{i,t}^*\le 5\mid x_{i,t},\alpha_i)+$$
$$+P(R_{i,t}=4\mid x_{i,t},\alpha_i^R)P(y_{i,t}^*\le 12.5\mid x_{i,t},\alpha_i)+$$
$$+P(R_{i,t}=5\mid x_{i,t},\alpha_i^R)P(y_{i,t}^*\le 25\mid x_{i,t},\alpha_i)]$$
$$\text{if } y_{i,t}=0$$

[7] The given likelihood contribution is for a respondent who participates and is asked the subjective probability question in each wave. $L_{i,t}^c$ is replaced by 1 if a respondent does not participate or is not asked the question in wave *t*.

$$L^c_{i,t}(\alpha_i,\alpha_i^R,\alpha_i^2,\alpha_i^3) = P(D_{i,t}=1\mid x_{i,t},\alpha_i^2,\alpha_i^3)[P(R_{i,t}=1\mid x_{i,t},\alpha_i^R)P(y^*_{i,t}>99.5\mid x_{i,t},\alpha_i)+$$
$$+P(R_{i,t}=2\mid x_{i,t},\alpha_i^R)P(y^*_{i,t}>97.5\mid x_{i,t},\alpha_i)+$$
$$+P(R_{i,t}=3\mid x_{i,t},\alpha_i^R)P(y^*_{i,t}>95\mid x_{i,t},\alpha_i)+$$
$$+P(R_{i,t}=4\mid x_{i,t},\alpha_i^R)P(y^*_{i,t}>87.5\mid x_{i,t},\alpha_i)+$$
$$+P(R_{i,t}=5\mid x_{i,t},\alpha_i^R)P(y^*_{i,t}>75\mid x_{i,t},\alpha_i)]$$
$$\text{if } y_{i,t}=100$$

$$L^c_{i,t}(\alpha_i,\alpha_i^R,\alpha_i^2,\alpha_i^3) = P(D_{i,t}=1\mid x_{i,t},\alpha_i^2,\alpha_i^3)[P(R_{i,t}=1\mid x_{i,t},\alpha_i^R)P(49.5<y^*_{i,t}\le 50.5\mid x_{i,t},\alpha_i)+$$
$$+P(R_{i,t}=2\mid x_{i,t},\alpha_i^R)P(47.5<y^*_{i,t}\le 52.5\mid x_{i,t},\alpha_i)+$$
$$+P(R_{i,t}=3\mid x_{i,t},\alpha_i^R)P(45<y^*_{i,t}\le 55\mid x_{i,t},\alpha_i)+$$
$$+P(R_{i,t}=4\mid x_{i,t},\alpha_i^R)P(37.5<y^*_{i,t}\le 62.5\mid x_{i,t},\alpha_i)+$$
$$+P(R_{i,t}=5\mid x_{i,t},\alpha_i^R)P(25<y^*_{i,t}\le 75\mid x_{i,t},\alpha_i)]$$
$$+P(D_{i,t}=3\mid x_{i,t},\alpha_i^2,\alpha_i^3)$$
$$\text{if } y_{i,t}=50$$

All probabilities in these expressions are either univariate cumulative normal probabilities or multinomial logit probabilities, and are therefore easy to compute.

The unconditional likelihood contribution of respondent *i* is the expected value of the conditional likelihood over the unobserved heterogeneity terms. Since $(\alpha_i,\alpha_i^R,\alpha_i^2,\alpha_i^3)$ can be written as Λu where u is a vector of six independent random variables with a standard normal distribution, we can also write the unconditional likelihood contribution as:

$$L_i = \int_{\mathbb{R}^4} L_i^c(\Lambda u)\varphi(u_1)...\varphi(u_4)\mathrm{d}u\,.$$

Here φ denotes the density of the standard normal distribution. To avoid numerical integration in five dimensions, the integral is replaced by a simulated mean:

$$SL_i = \frac{1}{M}\sum_{r=1}^{M} L_i^c(\Lambda u^r)\,,$$

where $u^1,...,u^M$ are simulated vectors with components drawn from independent standard normal distributions. To reduce the variance induced by the simulations, we used Halton draws to generate the u^r (see, e.g., Train, 2003). If M tends to ∞ at a fast enough rate, the

simulated maximum likelihood estimator is asymptotically equivalent to exact maximum likelihood (see, e.g., Gouriéroux and Monfort, 1996, or Hajivassiliou and Ruud, 1994).[8]

The estimated model parameters can be used to disentangle genuine 0, 50 and 100 answers from rounded and focal point answers. For example, the observed zeros can be genuine zeros that arise because of censoring a negative value of $y_{i,t}^*$, or they can be rounded zeros, e.g. if $y_{i,t}^* < 12.5$ and there is rounding to multiples of 25. The observed 50-s can be almost exact 50-s (if $R_{i,t} = 1$), they can be rounded 50-s (e.g., if $R_{i,t} = 5 \text{ and } 25 < y_{i,t}^* \leq 75$, but also if $R_{i,t} = 2$ and $47.5 < y_{i,t}^* \leq 52.5$, etc.), or they can be "50/50" answers that have no relation with $y_{i,t}^*$ (if $D_{i,t} = 3$). Once the model is estimated we can predict the probabilities of all of these types of responses, which will (if the model is correctly specified) add up to the observed fraction of 50 percent answers in the data.

Identification of the rounding probabilities intuitively relies on the sizes of the peaks in the observed frequency distribution. For example, the fact that we observe few probabilities that are not multiples of 5 (for a group with certain characteristics) implies that (for that group) the probability that $R_{i,t} = 1$ is small; the fact that we observe more multiples of 10 than other multiples of 5 determines the relative magnitudes of the probabilities that $R_{i,t} = 2$ and $R_{i,t} = 3$, etc.; the probability that $R_{i,t} = 5$ is determined by the difference between the fraction of zeros and 100-s in the data and the fraction implied by a censored regression model and the other types of rounding. The probability of a focal point 50/50 ($D_{i,t} = 3$) is determined by the difference between the observed fraction of 50-s and the fraction of 50-s predicted by a model with rounding only and no "50/50" answers.

4 Data and Description of Variables

We use data from the Health and Retirement Study (HRS) for 1994-2006.[9] The HRS is a biennial survey that started in 1992 with a national sample of 7,600 households with at least

[8] We found that the results were robust to the number and nature of the Halton draws. The results in the paper are based upon 20 Halton draws for each respondent.

one individual born between 1931 and 1941. Blacks, Hispanics, and Florida residents are over-sampled. Both the age-eligible respondents and their spouses are interviewed. The HRS collects information in a variety of areas, including socio-economic status, demographics, health, and family structure.[10] We use the original HRS cohort born between 1931 and 1941 and one additional birth cohort added in 1998, the so-called War Babies (WB), born between 1942 and 1947, and their spouses. We restrict out sample to individuals who are at least 50 years old. Tables 1 and 2 give an explanation of the variables used in the estimation and summary statistics for the estimation sample.

4.1 Subjective Probabilities

Respondents are asked a variety of subjective probability questions. At the beginning of the section in which these questions are asked, from 1996 on, respondents are given the following general explanation of the question type:

> *"Next I have some questions about how likely you think various events might be. When I ask a question I'd like for you to give me a number from 0 to 100, where "0" means that you think there is absolutely no chance, and "100" means that you think the event is absolutely sure to happen. "*

From 1998 on, this general explanation is followed by this example:

> *"For example, no one can ever be sure about tomorrow's weather, but if you think that rain is very unlikely tomorrow, you might say that there is a 10 percent chance of rain. If you think there is a very good chance that it will rain tomorrow, you might say that there is an 80 percent chance of rain."*

Immediately following this explanation (before the other probability questions) respondents in 2000 and 2002 were given what has become known as the sunshine question:

> *"Let's try an example together and start with the weather. What do you think are the chances that it will be sunny tomorrow? ("100" means a '100 percent chance of sunny weather'. And you can say any number from 0 to 100.)"*

In 1994, the wording of the explanation was slightly different:

> *"Now I am going to ask you about the chance of various events happening to you. Please answer the questions in terms of percent chance. Percent chance must be a*

---

[9] The HRS is sponsored by the National Institute of Aging (NIA) and conducted by the University of Michigan. We use the public use data files produced by the RAND Center for the Study of Aging (RAND HRS Data, Version I, and enhanced fat files).

[10] See Juster and Suzman (1995) and the HRS website at http://hrsonline.isr.umich.edu for an overview.

number from 0 to 100, where "0" means there is absolutely no chance, and "100" means that it is absolutely certain."

After this, respondents were given a slightly different version of the sunshine question:

"Let's start with the weather. Weather forecasters often say something like, "There's a 10-20 percent chance of rain tomorrow," meaning there's not much chance that it will rain. Using the same idea, what do you think the chances are that it will be sunny tomorrow?"

Some of the probability questions vary from wave to wave, and sometimes the wording differs. We conduct our analysis on answers to four subjective probability questions that were asked in most or all years for the HRS cohorts we are using and for which the wording is (almost) unchanged.

Two questions concern financial issues: bequests and inheritances. In 1994 and 1996, the question about bequests is as follows:[11]

"And what are the chances that you (or your husband/ wife/ partner) will leave an inheritance totaling $10,000 or more?"

From 1998 onwards, the wording of the question is somewhat different:

"Including property and other valuables that you might own, what are the chances that you (and your husband/ wife/ partner) will leave an inheritance totaling $10,000 or more?"

The wording of the second question is the same in all waves except for 2006:

"And how about the chances that you will receive an inheritance within the next ten years?"

In 2006, the wording was changed to:

"(Not counting anything you might give or leave to each other,) [what/What] are the chances that you (or your [husband/ wife/ partner]) will receive an inheritance during the next 10 years?"

The other two questions we consider are work related. One question is about the probability of working full time after a certain age:[12]

[11] If the answer is more than zero, a follow-up question asks about the chances of leaving a bequest of $100,000 or more. We used this question in a robustness check (coding the answer as equal to the first question if the answer was zero, "don't know" or "refuse"); the results differed in non-surprising ways.

"What do you think are the chances that you will work full-time after you reach age 62?"

Except for the year 2006, this question was supposed to be asked only if the respondent's age was below 62 and if the respondent was working. For reasons of consistency, we set all answers of respondents who were not working to missing as well as the answers of those respondents aged 62 and higher.

The last question we consider is about health and work:

"What about the chances that your health will limit your work activity during the next 10 years?"

The wording of this question was the same across all waves until 2002, but the question was not asked in 2004 and 2006. Like the other work-related question, it was only asked if the respondent was currently working. Answers of respondents who already had a work limiting health condition as well as answers given by proxy respondents were set to missing.

In Table 1, we show response patterns to the four probability questions under consideration for three mutually exclusive answer categories: "don't know" / "refuse", 50%, or a numerical value other than 50% for the latest available waves. Overall, the fraction of "don't know" / "refuse" answers is between 1.07% and 4.32%. The percentage of 50% answers is much larger for the question about a work limiting health condition than for the other three questions, suggesting that focal 50/50 answers might be a problem with this question in particular.

Figures 1 and 2 show histograms of the numerical answers to the four questions (50% is included; "don't know" / "refuse" is not) for all waves combined. For all questions, there is clear heaping at multiples of 5, and even more at multiples of 10, 25, and 50. While heaping at 50 may have a different cause (50/50 focal point answers - $D_{i,t} = 3$ in the model of Section 3), heaping at multiples of 5, 10 and 25 must be due to rounding. The histogram for the question on work related health problems not only shows a large frequency at 50%, but also relatively low frequencies at 0% and 100%. This suggests that rounding to multiples of 50 is not the explanation for the large frequency at 50% - if it were, we would also expect larger

[12] There was a follow-up question about the probability of working full-time after age 65. As with the bequest question, we did not use this because not everybody was asked. There were also many cases where respondents between the ages 62 and 65 were erroneously not asked this question.

frequencies at 0 (because of rounding values below 25 to 0) and 100 (rounding values above 75 to 100).

4.2 Covariates

The variables that might influence item nonresponse or induce focal point answers or rounding follow from the five respondent steps discussed in Section 2. Specifically, we include the following socio-demographic variables that might influence each of the five steps through their effects on knowledge, preferences, and response scales (Cao and Hill, 2005; Bassett and Lumsdaine, 2001): age, gender, race, education, marital status, and health. We also include cognitive ability, which influences item nonresponse and focal point answers (Knäuper et al., 1997; Hurd et al., 1998; Bassett and Lumsdaine, 2001). Following the existing literature, we proxy cognitive ability with answers to immediate word recall questions (see, for example, Cao and Hill, 2005 or van Soest and Hurd, 2008). Beginning in 1996, respondents were given 10 words and asked immediately to recall these words. In 1994, 20 words were given. To make the two tests comparable and to account for possible learning between waves, we use the percentile of the fraction of the words recalled within a wave, rather than the fraction of words recalled.[13] Missing values (1.3%) were set to zero.[14]

Additional variables that may be specifically important for some or all of the probability questions under consideration include:

- Income and wealth[15] (for all four questions we consider – these variables directly influence bequest possibilities; are probably correlated with wealth to be inherited; influence the value of continuing work; and may be correlated with health, type of work, job characteristics, and perceptions of which health conditions limit work);
- The effort and stress involved in the current job (for the work-related questions);

[13] We thank Patty St. Clair for bringing this pattern of the data to our attention.

[14] Missing observations in the cognition variable are related to cognition if individuals with difficulties in this area are less willing to answer this set of questions. We found evidence for this when comparing the results from probit estimations for a missing answer and OLS regressions on the word recall ratio or percentile with the same covariates - many of the coefficients have similar effects and significance levels in the two equations. We also experimented with an additional dummy for a missing word recall value, but this was never significant.

[15] We use the RAND model-based imputations for missing observations of income and wealth.

- Whether the respondent is the financial respondent of the household (for the bequest and inheritance questions, since this might influence knowledge about financial issues);
- Whether the respondent has health insurance (for the work related questions, since it increases the value of continued work, because health insurance often comes with the job);
- Whether the respondent has children[16] (for all questions, since many bequests are left to own children, family size is correlated across generations and influences the probability of inheritance, and the potential presence of grandchildren influences the value of leisure); and
- Whether the respondent's parents are alive (for all questions, since it influences the possibility of receiving an inheritance and may, therefore, affect wealth and the probability to leave a bequest; it may also determine the likelihood of withdrawing from the labor market in order to provide informal care (see, e.g., Heitmueller 2007).[17]

Table 2 describes the variables used in the analysis. We use observations on 14,584 individuals after dropping proxy respondents and those who answered only in one wave, those with missing information on education or race (30), and those who never answered any of the four probability questions (19). We dropped the individual observations by wave if information on health (36), age (1), marital status (78), or wealth (3) was missing. This resulted in 72,897 observations by individual by wave.

Table 3 shows the descriptive statistics for the whole sample and by work status. On average, the individuals in our sample are 62 years old; they are more likely to be female, and most of them are married and have children. Almost one third of the respondents had at least one living parent.

5 Estimation Results

Before discussing the results of our main model specifications, we first present some simulation results for estimated model specifications without any covariates, to give a first

[16] We also experimented with the number of children but this did not improve the results. The variable for whether the respondent has any children is set to zero if missing (between 0 and 234 observations per wave).

[17] This variable is set to zero if missing (between 187 and 297 observations per wave).

impression of the importance of rounding and focal point 50/50 answers.[18] Table 4 shows the simulated probabilities for each possible type of response for each of the four expectation questions for the latest available wave (they are very similar to those for the other waves). For the inheritance and bequest expectations, the majority of 50 percent answers are related to rounding and not to uncertainty leading to a focal 50/50 answer: the estimated probabilities of a focal 50/50 are very small. This probability is somewhat larger for the question on working until age 62 and quite substantial for the question on work related health. This corresponds to what we saw in Figure 2: the work related health question has many more 50 percent-answers and relatively few 0 or 100 percent-answers, and this distribution is hard to explain with rounding only.

The opposite pattern is observed for extreme rounding: rounding to the nearest multiple of 50 is common for the inheritance and bequest questions, less common for the questions on working after age 62, and does not occur at all for the question on work related health. On the other hand, rounding to multiples of 10 is the most common form of rounding for all questions, corresponding to the fact that heaping at multiples of 10 is a salient feature of all the histograms in Figures 1 and 2.

The small probabilities of focal 50/50 answers in the model without regressors suggest that estimating a full model with all regressors may not be feasible in all cases. In particular, it may well be the case that the estimated probability of a focal 50/50 answer tends to zero for specific socio-economic groups, implying a coefficient of minus infinity for a dummy for that group. This is indeed what we find in some cases, particularly for the probability of leaving a bequest of at least USD 10,000, which had the smallest 50/50 probability in Table 4. In these cases, some specific coefficients converge to minus or plus infinity, implying a zero probability for the specific socio-economic group (see Table 8 below for details).

In what follows, we present the estimation and simulation results for the complete models. We begin by discussing the simulated probabilities, followed by the determinants of rounding, nonresponse and 50/50 answers. We then present the estimation results of the subjective unobserved probabilities (equation 1) and compare them to the results of standard random-effects tobit models accounting for the censoring at 0 and at 100. We conclude by presenting the covariance matrices of the individual effects.

[18] Estimation results for these specifications are available upon request.

Table 5 shows the simulated probabilities for the various response types for each of the four expectation questions for the latest available wave, similar to Table 4. They are qualitatively similar to the simulated probabilities for the other waves, although the differences are larger than in Table 4, mainly due to the time dummies, which are sometimes significant. The qualitative results are very much in line with those in Table 4. With the exception of work limitations, the large majority of 50 percent-answers are related to rounding and not to focal 50/50 answers reflecting complete uncertainty. For all questions, rounding to the nearest multiple of 10 is most frequent, followed by rounding to multiples of 25.

The simulated probabilities for nonresponse are very close to the actual frequencies shown in Table 3. Of the 50 percent-answers in Table 3, only a limited part is explained by focal point 50/50 answers; the remainder is explained by rounding to multiples of 5, 10, 25, or 50. The inheritance question shows less rounding to multiples of 10 or 25 than the other questions. The response patterns for the work limitations question deviates from the other three in the sense that there is a substantial number of focal 50/50 answers but no rounding to multiples of 50, as we saw in Table 4. On the other hand, the chances of rounding to multiples of 5, 10, or 25 are of the same order of magnitude as for the other questions.

The estimation results for the rounding equations can be seen in Table 6. Many coefficients are similar across the four questions. There is less rounding the higher the education and the better cognition, corresponding to the notion that the more skilled give more precise answers. Blacks are less likely to round than whites, but the differences between Hispanics and whites vary across questions. Higher income is related to less rounding for the bequest questions, probably because of lower uncertainty. Age is positively related to rounding, except for the inheritance question, which is probably because uncertainty about receiving an inheritance falls with age.

Tables 7 and 8 present the results of the multinomial logit regression explaining the choice between giving a (rounded) numerical answer (the benchmark), answering “don’t know” or “refuse” (nonresponse), or giving a focal 50/50 answer. We find that higher education, higher income, and better cognition are related to a lower probability of nonresponse. Blacks and Hispanics are more likely to not respond than whites. Financial respondents are more likely to

respond to the expectation questions about inheritance and bequests.[19] The nonresponse probability increases significantly with age for three out of the four questions. The effect of wealth is inconsistent across questions – wealthier households significantly less often respond to the bequest question, but more often to the question on expected work limiting health problems.

Table 8 shows the results for the tendency to give a focal 50/50 answer in the multinomial logit model. Here the pattern is less consistent across questions than in the previous table. For example, age significantly increases the focal answer probability for the work related health problems question, but reduces it for the probability questions on working full time after age 62 and receiving an inheritance. The estimated coefficient of minus infinity for Hispanics for the inheritance question implies that Hispanics, the estimated probability of a focal point answer is zero. Similarly, we found a zero probability of a focal 50/50 for families without children (*child=0*) for the bequest question. There is no significant effect of our measure of cognition, and the effect of education is not consistent across questions. For the work limitations question, we find a strong and significant negative effect of education. Compared to otherwise similar respondents with high school education, respondents with a college degree are about 8.6 percentage points less likely to give a focal 50/50 answer.[20] For the other questions the focal point probabilities seem to be too small to get reliable estimates of their determinants (see also the discussion of Table 4 above).

Tables 9 and 10 show the estimation results of the unobserved probabilities (equation (1) in Section 3) and compare them to estimates from random-effects tobit estimations that do not account for rounding or focal 50/50 answers and assume item nonresponse is random conditional upon the covariates. In general, the results of the estimates using the full model are qualitatively similar to those of the random-effects (RE) tobit models. Signs, magnitudes, and significant levels of the parameter estimates usually do not differ much, though there are a few exceptions. This seems reassuring for researchers who take the reported probabilities

[19] This is similar to the findings of Cao and Hill (2005) who analyze item nonresponse in questions about assets in the 2000 wave of the HRS.

[20] Using the parameter estimates in Tables 7 and 8, the estimated marginal effect of changing from high school to college education, keeping other variables constant, is given by -0.524 P(50/50) (1-P(50/50) + 0.754 P(50/50) P(nonresponse). For the average probabilities of a focal 50/50 and item nonresponse in Table 5, this gives -0.524*0.221*(1-0.221)+0.754*0.221*0.045=-0.083.

for granted. Our findings imply that not taking account of rounding, focal 50/50 answers, or item nonresponse, does not have a large impact on the conclusions concerning the determinants of the subjective probabilities.

We find that higher age and bad health strongly increase the probability of a work limiting health condition, as does work that is stressful or that requires physical effort (Table 9). Blacks have lower probabilities than whites. According to the full model results, Hispanics and high income groups have significantly lower probabilities of work limiting health problems than their white and low income counterparts, while these effects are not significant in the RE tobit. For most covariates, the effects are somewhat larger according to the complete model based estimates than in the RE tobit specifications.

The probability of working past age 62 increases with age, higher education and cognition, and falls significantly with income and wealth (Table 9). Married women are much less likely to work past age 62 than married men or unmarried men or women. Having health insurance is positively related to the probability of working past age 62. This is likely to be related to employer provided health insurance, which makes it unattractive to stop working before age 65, the age of Medicare eligibility. Having a stressful job increases the probability to work after age 62 and the effect is significant. An explanation for this counterintuitive result might be correlated unobserved factors that make someone select a stressful job and increase the motivation for continued work; such factors are not explicitly taken into account in the models.

The patterns of the covariates for the probabilities of receiving an inheritance and leaving bequests are clear and strong (shown in Table 10). As expected, the variables related to socio-economic status - education, marriage, income, and wealth - are strongly positively related to these two probabilities. Even after controlling for income and wealth, blacks and Hispanics report much lower probabilities than whites and non-Hispanics. The main difference between RE tobit and the complete model is in the *work* variable: workers have significantly higher chances of receiving an inheritance or leaving a bequest of at least USD 10,000 according to the RE tobit estimates, but the effect is small and insignificant in the estimates based upon the complete model. In addition, the RE tobit estimates of the effects of basic demographics are always larger than those in the complete model, suggesting that not explicitly modeling the rounding and the response decisions leads to an overestimation of the effects of education, gender, race, and ethnicity.

For the probability of receiving an inheritance, having at least one living parent is the most important variable. Age and bad health are negatively related to the probability of receiving an inheritance; cognition has a positive effect. A possible interpretation is that these are all indicators of mortality expectations, and that people who expect to live longer also have higher chances of receiving an inheritance.

With respect to bequest probabilities, men report higher probabilities, which could be related to their lower life expectancies. Individuals in bad health report lower probabilities, possibly related to higher expected out-of-pocket medical expenses. Individuals with children report higher bequest probabilities, likely because they have a stronger bequest motive.

Lastly, Table 11 shows the covariance matrices of the individual effects for the four probability questions. For all four probabilities, there is significant unobserved heterogeneity in all parts of the model: the latent unobserved probabilities (equation (1)), the tendency of rounding, the probability of item non-response, and the probability of a focal 50/50 answer. The random effects in the unobserved heterogeneity are of similar magnitude as would be implied by a random effects tobit model (see the bottom panels of Tables 9 and 10). Moreover, the correlations between the various unobserved heterogeneity terms are usually significant as well. For example, respondents who tend to not respond also tend to give a focal 50/50 answer: the correlation coefficient between the random effects is always significantly positive, varying from 0.48 (for *worklm*) to 0.97 (for *beq10)* (not shown but implied by Table 11). This is in line with the notion that both item non-response and focal fifty-fifties are an expression of persistent uncertainty over time. The correlation between unobserved heterogeneity terms in the rounding equation and the item non-response tendency is also positive in three out of four cases, suggesting that rounding is related to uncertainty but also to other factors (like cultural differences, cf., e.g., the finding that blacks are less likely to give a rounded answer than whites, while they are more likely to answer "don't know" or "refuse").

6 Conclusions

Answers to expectation questions are used more and more frequently to understand individual behavior and outcomes. How individuals respond to this type of questions is not well understood. Answers show much heaping at specific probabilities, which is often interpreted as individuals giving focal point answers and taken as evidence that these answers do not reflect the actual subjective probabilities. Heaping, however, could also be the result of

rounding. In this paper, we develop a random-effects panel data model of response to questions about subjective probabilities, explicitly accounting for the possibility of focal point answers of 50, rounding, and item nonresponse. Disentangling the contribution to the response of each of these three types of answers allows an assessment of the validity of the response as well as of the importance of taking these into account when using the responses to subjective probability questions to explain individual behavior or outcomes.

We find that most answers of 50 are related to rounding and not to focal point answers. Rounding and nonresponse are strongly related to socio-demographic variables, including race and ethnicity. When comparing the results from our model with an unconditional random-effects tobit model, we find that the coefficients are generally very similar, though there are some exceptions. Not taking into account the possibility of rounding, focal points answers and item nonresponse leads to an overestimation of the effects of socio-demographic variables on the expected probabilities to receive an inheritance or leave a bequest. This also implies that, for example, blacks will be underrepresented if data with item nonresponse is discarded. It also means that inferences from answers to expectation questions might be different depending on the group analyzed, because the same answer category might have different meanings for different groups. Further research is needed as to why response behavior differs among the groups, and how the possible answer categories – such as 50/50 - are interpreted by the respondents.

References

Bassett, W. F. and R. L. Lumsdaine (2001). Probability limits. *Journal of Human Resources* 36(2), 327-363.

Beatty, P. and D. Hermann (2002). To answer or not to answer: Decision processes related to survey item nonresponse. In: R. M. Groves et al. (eds.), *Survey Nonresponse*. John Wiley, New York, 71-85.

Bernheim, B. D. (1990). How do the elderly form expectations? An analysis of responses to new information. In: D. A. Wise (ed.): *Issues in the Economics of Aging*. Chicago, NBER-University of Chicago Press, 259-283.

Brown, J. R., C. C. Coile and S. J. Weisbenner (2010). The effect of inheritance receipt on retirement. *Review of Economics and Statistics* 92(2), 425-434.

Bruine de Bruin, W. B. Fischhoff, S. G. Millstein, and B. L. Halpern-Felsher (2000). Verbal and numerical expressions of probability: It's a fifty-fifty chance. *Organizational Behavior and Human Decision Processes* 81(1), 115-131.

Cao, H. and D. H. Hill (2005). Knowledge and preference in reporting financial information. *Michigan Retirement Research Center Working Paper* 2005-100.

Cramer, A. T. and G. A. Jensen (2006). Why don't people buy long-term-care insurance? *Journal of Gerontology: Social Sciences* 61B(4), S185-S193.

Dominitz, J. and C. F. Manski (2006). Measuring pension benefit expectations probabilistically. *Labour* 20(2), 201-236.

Dustmann, C. and A. van Soest (2001). Language fluency and earnings: estimation with misclassified language indicators. *Review of Economics and Statistics* 83(4), 663-674.

Fox, C. R. and A. Tversky (1995). Ambiguity aversion and comparative ignorance. *Quarterly Journal of Economics* 110(3), 585-603.

Gouriéroux, C. and A. Monfort (1996). *Simulation-based Econometric Methods*. Oxford University Press, Oxford.

Hajivassiliou, V. and P. Ruud (1994). Classical estimation methods for LDV models using simulation. In: R. Engle and D. McFadden (eds.), *Handbook of Econometrics, vol. IV*, North-Holland, New York, 2384-2443.

Hausman, J. A., Abrevaya, J., and Scott-Morton, F. M. (1998). Misclassification of a dependent variable in a discrete response setting. *Journal of Econometrics* 87, 239–269.

Heitjan, D.F. and D.B. Rubin (1990). Inference from coarse data via multiple imputation with application to age heaping, *Journal of the American Statistical Association* 85(410), 304-314.

Heitmueller, A. (2007). The chicken or the egg? endogeneity in labour market participation of informal carers in England. *Journal of Health Economics* 26, 536-559.

Hurd, M., D. McFadden, and L. Gan (1998). Subjective survival curves and life cycle behavior. In: D. A. Wise (ed.), *Inquiries in the Economics of Aging*, Chicago, University of Chicago Press, 295-305.

Hurd, M. D. and J. P. Smith (2002). Expected bequests and their distribution. NBER Working Paper 9142.

Juster, F. T. and R. Suzman (1995). An overview over the Health and Retirement Study. *Journal of Human Resources* 30 (Supplement), S7-S56.

Kapteyn, A., K. J. Kleinjans and A. van Soest (2009). Intertemporal consumption with directly measured welfare functions and subjective expectations. *Journal of Economic Behavior & Organization* 72, 425-437.

Kézdi, G. and R. J. Willis (2009). Stock market expectations and portfolio choice of American households. Mimeo, University of Michigan.

Knäuper, B., R. F. Belli, D. H. Hill, and A. R. Herzog (1997). Question difficulty and respondents' cognitive ability: The effect on data quality. *Journal of Official Statistics* 13(2), 181-199.

Lillard, L. and R. J. Willis (2001). Cognition and wealth: The importance of probabilistic thinking. University of Michigan Retirement Research Center Working Paper WP 2001-001.

Maestas, N. (2010). Back to work: expectations and realizations of work after retirement. *Journal of Human Resources* 45(3), 718-747.

Manski, C. F. (2004). Measuring expectations. *Econometrica* 72(5), 1329-1376.

Manski, C. F. and F. Molinari (2010). Rounding probabilistic expectations in surveys. *Journal of Business and Economic Statistics* 28(2), 219-231.

Michaud, P.-C. and S. Rohwedder (2008). Forecasting labor force participation and economic resources of the early baby boomers. University of Michigan Retirement Research Center Working Paper WP 2008-175.

Michaud, P.-C. and A. van Soest (2008). How did the elimination of the earnings test above the normal retirement age affect labour supply expectations? *Fiscal Studies* 29(2), 197-231.

Mok, S. (2010). Precautionary savings or bequest motives? Evidence from the health spending of the elderly. Mimeo, University of Wisconsin.

Quandt, R. E., and J.B. Ramsey (1978). Estimating mixtures of normal distributions and switching regressions, *Journal of the American Statistical Association*, 73, 730-738.

Scholz, J. K., A. Seshadri and S. Khitatrakun (2006). Are Americans savings "optimally" for retirement? *Journal of Political Economy* 114(4), 607-643.

Schwarz, N. and D. Oyserman (2001). Asking questions about behavior: Cognition, communication, and questionnaire construction. *American Journal of Evaluation* 22(2), 127-160.

Shoemaker, P. J., M. Eichholz, and E. A. Skewes (2002). Item nonresponse: Distinguishing between don't know and refuse. *International Journal of Public Opinion Research* 14(2), 193-201.

Tourangeau, R., L. J. Rips, and K. Rasinski (2000). *The psychology of survey response*. Cambridge, UK, Cambridge University Press.

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

Van Soest, A. and M. D. Hurd (2008). A test for anchoring and yea-saying in experimental consumption data, *Journal of the American Statistical Association* 103(481), 126-136.

Table 1 Response Patterns by Probability Question for Latest Available Wave*

Question	Don't know / Refuse	50%	Numerical response other than 50%	Total
Beq10	4.27%	6.69%	89.04%	100%
Inher	3.24%	3.90%	92.86%	100%
Work62	1.07%	14.74%	84.19%	100%
Worklm	4.32%	31.81%	63.86%	100%

** 2002 for work limitations, 2006 otherwise*

The response patterns of the bequest and inheritance questions are for the full sample, the work-related questions for working respondents only.

Figure 1 Distribution of Responses: Inheritances and Bequests

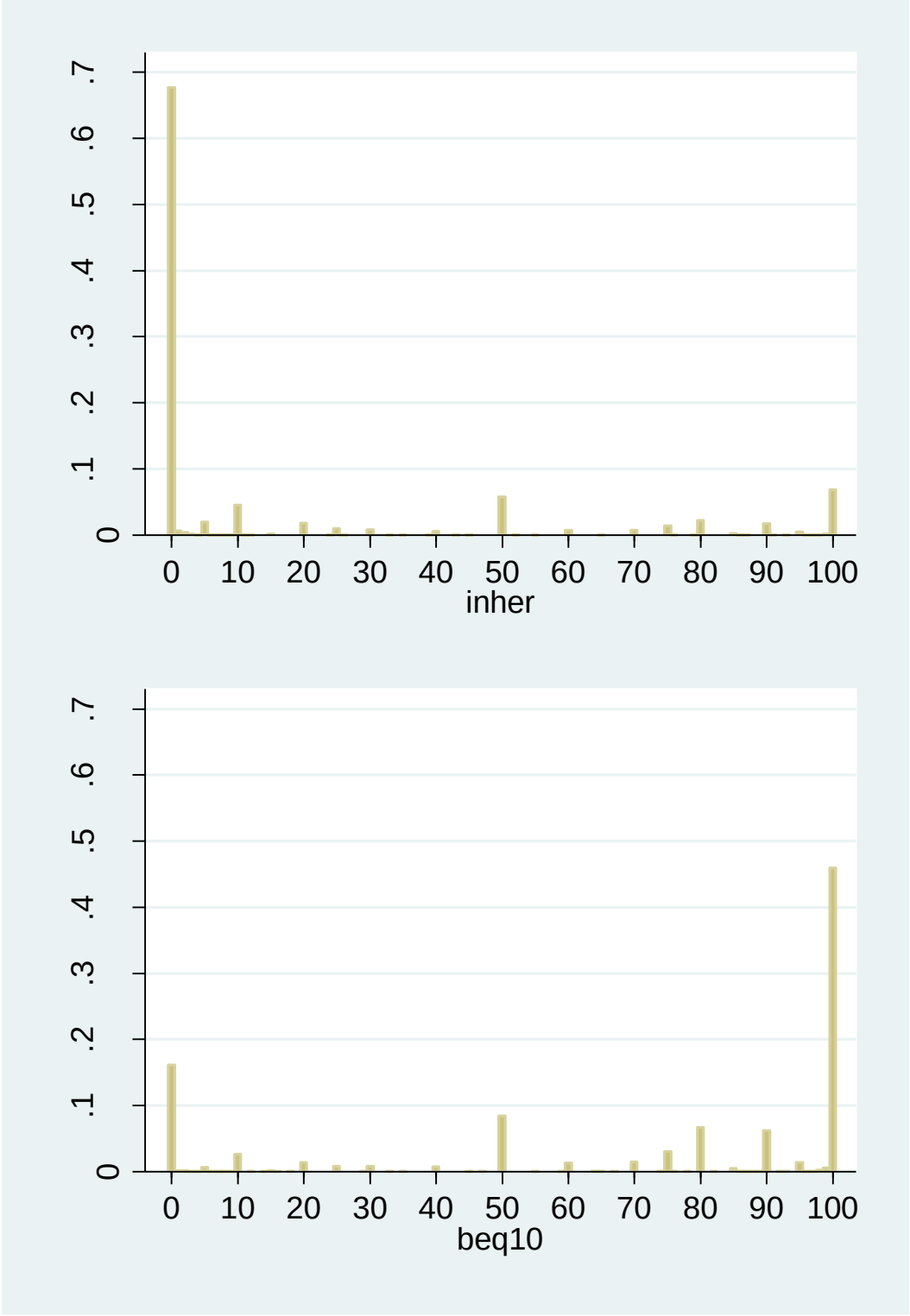

Figure 2 Distribution of Responses: Work-Related Questions

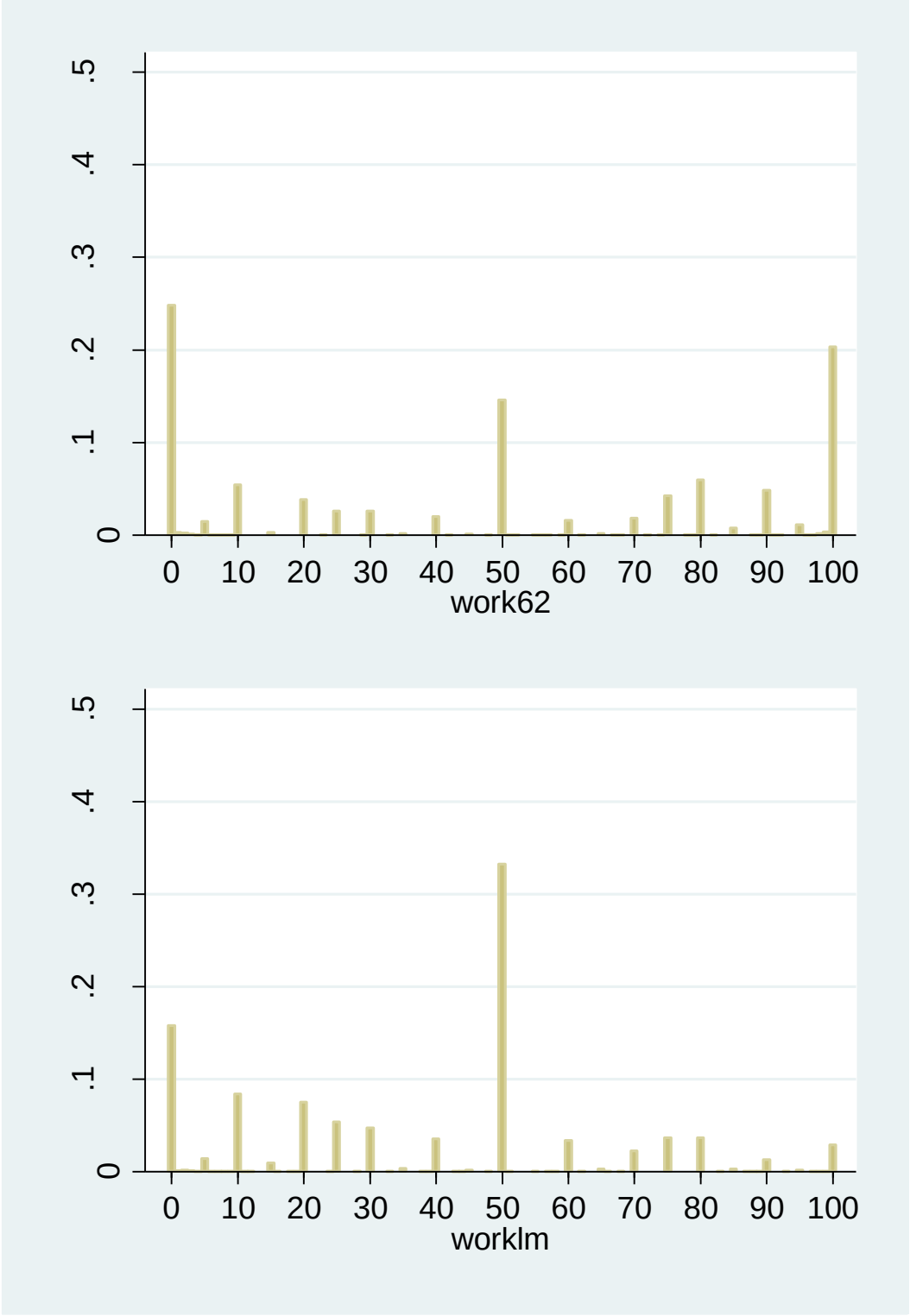

Table 2 Variable Definitions

Variable Name	Description
Age	Age in years
Male	= 1 if male
Black	= 1 if Black
Hispanic	= 1 if Hispanic
Education	
Less than HS	= 1 if less than High School
HS	= 1 if High School or GED
Some College	= 1 if Some College
College	= 1 if College or more
Not Married	= 1 if neither married nor partnered
Child	= 1 if at least one child
Parent alive	= 1 if any of the parents is alive
Financial Respondent	= 1 if financial respondent
Bad Health	= 1 if self-reported health status fair or poor
Cognition ratio	Ratio of words recalled in immediate recall questions
Cognition	Percentile of Cognition ratio divided by 100
Income	Real total household income in $2002 divided by the square root of the household size
Log of income	Log of real income defined as log(income+1) if income >=0 and –log(1-income) if income<0
Wealth	Real total household wealth in $2002 divided by the square root of the household size
Log of wealth	Log of real wealth defined as log(wealth+1) if wealth >=0 and –log(1-wealth) if wealth<0
Insured	= 1 if covered by health insurance (by government, current or previous employer, other)
Work	=1 if currently working for pay
Effort	= 1 if current job requires lots of physical effort or lifting heavy loads or stooping, kneeling, or crouching all, almost all, or most of the time
Stress	= 1 if current job involves a lot of stress all, almost all, or most of the time
Probability Questions	
Inher	Probability of receiving an inheritance within next 10 years
Beq10	Probability of leaving a bequest of USD 10,000 or more
Work62	Probability of working past age 62
Worklm	Probability of work limiting health condition during next 10 years

Table 3 Summary Statistics (all years)

Variable	All	Non-Working	Working
Age	61.89	64.06	59.66
	(6.19)	(6.10)	(5.44)
Male	0.4339	0.3977	0.4710
Black	0.1529	0.1649	0.1403
Hispanic	0.0842	0.0986	0.0694
Education			
Less than HS	0.2243	0.2897	0.1572
HS	0.3731	0.3825	0.3636
Some College	0.2067	0.1829	0.2310
College	0.1960	0.1449	0.2482
Not Married	0.2555	0.2739	0.2365
Child	0.9286	0.9297	0.9274
Parent alive	0.3066	0.2324	0.3828
Financial Respondent	0.6611	0.6514	0.6708
Bad Health	0.2468	0.3570	0.1338
Cognition Ratio	0.5500	0.5258	0.5749
	(0.187)	(0.188)	(0.183)
Log of Income	10.22	9.850	10.59
	(1.379)	(1.520)	(1.096)
Log of Wealth	10.52	10.14	10.91
	(4.479)	(4.862)	(4.013)
Work	0.4938	0	1
Insured	0.7841	0.7956	0.7724
Effort			0.4023
Stress			0.5347
N	72,897	36,885	35,987

Standard deviations shown in parentheses except for dummy variables.

Table 4 Simulated Probabilities for Latest Available Wave*:

Model without Regressors

	Work-related Questions		Inheritances and bequests	
	Worklm	**Work62**	**Inher**	**Beq10**
Rounding				
To a multiple of				
1	1.27	3.04	10.35	4.85
5	11.58	16.53	25.96	15.05
10	43.51	49.40	38.26	45.72
25	16.65	19.00	8.58	12.09
50	0.00	4.80	13.78	18.95
Nonresponse	4.62	1.04	2.62	3.33
50/50	22.37	6.19	0.45	0.02
N	4,278	1,309	10,301	10,301

** 2002 for work limitations, wave 2006 otherwise.*

Table 5 Simulated Probabilities for Latest Available Wave*: Complete Model

	Work-related Questions		Inheritances and bequests	
	Worklm	**Work62**	**Inher**	**Beq10**
Rounding				
To a multiple of				
1	1.349	3.74	12.45	5.30
5	11.88	17.97	28.56	16.35
10	43.93	46.47	38.19	46.39
25	16.25	15.67	7.895	11.26
50	0.00	10.56	8.92	16.09
Nonresponse	4.50	1.07	3.44	4.56
50/50	22.11	4.54	0.54	0.05
N	4,278	1,309	10,301	10,301

** 2002 for work limitations, wave 2006 otherwise.*

Table 6 Estimation Results for Rounding Equations*

	Work-related questions		Inheritances and bequests	
	Worklm	**Work62**	**Inher**	**Beq10**
Age/100	1.316	3.085	-0.981	0.900
	(4.11)	(5.23)	(-2.96)	(4.20)
Male	0.001	-0.090	-0.117	-0.122
	(0.03)	(-2.18)	(-3.02)	(-4.60)
Black	-0.333	-0.097	-0.299	-0.049
	(-7.18)	(-1.95)	(-5.84)	(-1.65)
Hispanic	-0.275	-0.082	-0.032	0.286
	(-3.65)	(-1.15)	(-0.44)	(7.16)
Education				
Less than high school	0.016	0.187	0.116	0.222
	(0.34)	(3.45)	(2.22)	(7.69)
Some college	-0.068	-0.123	-0.026	-0.183
	(-1.73)	(-2.71)	(-0.60)	(-6.04)
College	-0.196	-0.385	-0.178	-0.585
	(-5.10)	(-8.03)	(-4.05)	(-17.82)
Not married	0.040	0.053	0.003	0.039
	(0.91)	(1.00)	(0.06)	(1.18)
Not married * male	-0.067	0.010	0.001	0.124
	(-0.91)	(0.11)	(0.01)	(2.41)
Child	0.038	0.035	0.051	-0.017
	(0.66)	(0.53)	(0.80)	(-0.42)
Parent alive	-0.031	0.053	0.181	-0.048
	(-1.04)	(1.51)	(5.34)	(-2.03)
Financial respondent	-	-	-0.002	-0.014
			(-0.05)	(-0.56)
Bad health	0.031	0.037	0.051	0.075
	(0.63)	(0.72)	(1.23)	(3.16)
Cognition	-0.057	-0.164	-0.073	-0.187
	(-1.07)	(-2.68)	(-1.29)	(-4.88)
Log of income /100	-0.265	-0.956	-1.115	-5.202
	(-0.17)	(-0.56)	(-0.87)	(-6.84)
Log of wealth /100	0.390	0.128	1.422	0.059
	(0.97)	(0.30)	(3.63)	(0.27)
Work	-	-	0.008	0.000
			(0.74)	(-0.03)

** t-values shown in parentheses. Intercept and time dummies included but not reported.*

Table 7 Estimation Results for Random Effects Multinomial Logit: Nonresponse*

	Work-related questions		Inheritances and bequests	
	Worklm	**Work62**	**Inher**	**Beq10**
Age/100	4.504	2.076	2.545	4.251
	(5.00)	(0.81)	(4.96)	(9.20)
Male	-0.147	-0.402	-0.450	-0.409
	(-1.60)	(-2.29)	(-6.43)	(-6.33)
Black	0.580	0.760	0.967	0.805
	(5.88)	(4.46)	(14.60)	(12.24)
Hispanic	0.846	1.244	0.165	0.380
	(6.48)	(6.21)	(1.87)	(4.73)
Education				
Less than high school	0.563	0.511	0.574	0.571
	(5.24)	(2.92)	(8.37)	(9.06)
Some college	-0.071	-0.106	0.062	-0.208
	(-0.72)	(-0.54)	(0.76)	(-2.72)
College	-0.754	-0.395	0.028	-0.261
	(-6.28)	(-1.54)	(0.34)	(-3.24)
Not married	0.237	0.144	-0.062	-0.136
	(2.19)	(0.79)	(-0.74)	(-1.76)
Not married * male	-0.027	0.434	0.550	0.346
	(-0.15)	(1.45)	(4.43)	(2.92)
Child	0.111	-0.164	-0.242	-0.223
	(0.69)	(-0.63)	(-2.54)	(-2.42)
Parent	0.060	0.235	0.338	-0.063
	(0.75)	(1.71)	(5.30)	(-0.99)
Financial respondent	-	-	-0.257	-0.341
			(-3.82)	(-5.68)
Bad health	0.060	0.465	0.071	0.123
	(0.54)	(3.04)	(1.15)	(2.24)
Cognition	-0.352	-1.216	-1.468	-1.090
	(-2.66)	(-4.74)	(-14.33)	(-11.73)
Log of income/100	-11.311	-20.215	-12.768	-13.433
	(-4.00)	(-5.61)	(-8.12)	(-8.90)
Log of wealth/100	-2.400	0.316	-0.0721	1.698
	(-2.96)	(0.21)	(-0.12)	(2.91)
Work	-	-	-0.039	-0.042
			(-5.47)	(-6.78)

** t-values shown in parentheses. Intercept and wave dummies included but not reported.*

Table 8 Estimation Results for Random Effects Multinomial Logit: 50/50*

	Work-related questions		Inheritances and bequests	
	Worklm	**Work62**	**Inher**	**Beq10**
Age/100	3.571	-19.513	-11.421	-1.783
	(6.08)	(-5.11)	(-6.56)	(-0.15)
Male	-0.011	-0.016	-0.207	-3.029
	(-0.17)	(-0.08)	(-1.23)	(-1.27)
Black	-0.313	0.578	-0.324	-2.821
	(-3.83)	(2.41)	(-1.19)	(-1.40)
Hispanic	-0.122	0.402	-∞	-1.089
	(-1.14)	(1.25)		(-0.36)
Education				
Less than high school	0.017	-0.234	-0.216	0.076
	(0.20)	(-0.73)	(-0.83)	(0.04)
Some college	-0.173	-0.010	0.077	-6.088
	(-2.48)	(-0.04)	(0.49)	(-2.07)
College	-0.524	-0.189	-0.599	0.867
	(-7.12)	(-0.76)	(-2.91)	(0.37)
Not married	-0.044	-0.536	-0.333	0.000
	(-0.54)	(-1.65)	(-1.46)	(0.00)
Not married * male	-0.024	-0.463	0.092	0.000
	(-0.18)	(-0.70)	(0.25)	(0.00)
Child	-0.116	-0.364	-0.391	--**
	(-1.12)	(-1.30)	(-1.62)	
Parent alive	-0.152	-0.422	2.604	0.327
	(-2.79)	(-2.19)	(7.83)	(0.22)
Financial respondent	-	-	0.513	0.143
			(3.16)	(0.10)
Bad health	-0.142	-0.049	-0.595	5.377
	(-1.66)	(-0.18)	(-2.29)	(2.02)
Cognition	-0.112	-0.182	0.308	4.290
	(-1.20)	(-0.53)	(1.17)	(1.57)
Log of income/100	-1.586	-6.870	3.479	255.78
	(-0.61)	(-0.78)	(0.49)	(1.77)
Log of wealth/100	-0.503	-0.362	2.473	-44.305
	(-0.74)	(-0.17)	(1.15)	(-2.19)
Work	-	-	0.175	7.150
			(1.03)	(1.53)

** t-values in parentheses. Intercept and time dummies included but not reported.*

*** For child=0, the estimated probability converged to 0.*

Table 9 Results Main Equation in Full Model & Random Effects Tobits: Work-related probabilities (t-values in parentheses)

	Worklm		Work62	
	Full Model	**RE Tobit**	**Full Model**	**RE Tobit**
Age	0.906	0.776	1.381	1.452
	(13.33)	(13.60)	(9.77)	(8.85)
Male	2.013	1.607	16.910	19.042
	(2.77)	(2.62)	(13.82)	(13.09)
Black	-5.377	-4.694	-14.303	-15.083
	(-6.21)	(-5.98)	(-8.72)	(-8.02)
Hispanic	-2.747	-1.945	-1.300	0.493
	(-2.20)	(-1.82)	(-0.57)	(0.20)
Education				
Less than high school	-1.576	-0.909	-2.510	-3.425
	(-1.72)	(-1.10)	(-1.48)	(-1.70)
Some college	-0.608	-1.049	6.067	6.433
	(-0.73)	(-1.50)	(4.32)	(3.82)
College	0.556	-1.149	7.383	8.437
	(0.64)	(-1.61)	(5.09)	(4.81)
Not married	2.200	1.442	17.257	19.166
	(2.39)	(1.85)	(11.89)	(11.28)
Not married * male	-4.044	-2.411	-18.180	-19.891
	(-2.75)	(-1.88)	(-7.84)	(-7.03)
Child	1.967	1.008	0.521	0.680
	(1.74)	(1.05)	(0.29)	(0.33)
Parent alive	-2.395	-2.331	2.474	2.846
	(-3.92)	(-4.60)	(2.74)	(2.65)
Bad health	15.593	11.721	-10.339	-11.127
	(18.82)	(15.86)	(-9.10)	(-8.54)
Cognition	-0.437	-0.473	2.693	3.487
	(-0.45)	(-0.60)	(2.01)	(2.31)
Log of income	-0.295	-0.294	-0.967	-1.101
	(-1.09)	(-1.31)	(-2.88)	(-2.63)
Log of wealth	-0.141	-0.099	-0.635	-0.664
	(-2.02)	(-1.64)	(-6.60)	(-6.03)
Insured	0.660	0.810	3.056	2.974
	(1.02)	(1.52)	(3.41)	(2.85)
Effort	2.281	1.979	-0.232	-0.042
	(3.87)	(3.96)	(-0.27)	(-0.04)
Stress	2.547	1.998	3.305	3.636
	(4.72)	(4.50)	(4.31)	(4.17)
Sigma u	19.128	16.658	44.251	48.474
	(53.60)	(57.15)	(60.96)	(76.14)
Sigma e	28.120	26.591	36.500	42.214
	(146.49)	(154.69)	(99.60)	(122.06)
N		23,785**		23,346**

** Time dummies and intercept included but not reported.*

*** Excluding item nonresponse*

Table 10 Results Main Equation in Full Model & Random Effects Tobits: Probabilities on inheritances and bequests (t-values in parentheses)

	Inher		Beq10	
	Full Model	**RE Tobit**	**Full Model**	**RE Tobit**
Age	-2.125	-2.674	-0.168	-0.025
	(-24.86)	(-22.19)	(-2.63)	(-0.27)
Male	4.664	7.413	11.769	15.152
	(4.49)	(5.01)	(14.36)	(12.99)
Black	-14.105	-15.224	-19.126	-25.903
	(-10.63)	(-8.13)	(-19.29)	(-18.25)
Hispanic	-37.768	-46.733	-11.434	-15.868
	(-21.69)	(-17.12)	(-9.36)	(-8.53)
Education				
Less than high school	-21.671	-25.016	-22.387	-28.844
	(-16.99)	(-13.12)	(-23.91)	(-20.73)
Some college	15.536	16.140	11.760	14.427
	(13.19)	(9.54)	(12.43)	(10.60)
College	22.575	26.899	21.999	25.025
	(18.16)	(15.33)	(20.99)	(17.27)
Not married	-5.660	-7.196	-8.192	-10.222
	(-4.67)	(-4.19)	(-9.23)	(-7.90)
Not married * male	-3.480	-3.945	3.190	4.857
	(-1.80)	(-1.50)	(2.45)	(2.43)
Child	-0.846	-3.177	7.586	9.424
	(-0.54)	(-1.58)	(7.14)	(5.95)
Parent alive	46.736	54.097	0.874	0.629
	(64.11)	(52.47)	(1.36)	(0.74)
Financial respondent	-1.499	-1.262	0.842	1.114
	(-1.60)	(-0.95)	(1.13)	(1.07)
Bad health	-6.958	-6.906	-10.771	-12.218
	(-8.00)	(-6.41)	(-18.91)	(-15.39)
Cognition	4.433	4.796	11.828	14.002
	(3.68)	(3.47)	(13.29)	(12.37)
Log of income	1.830	1.690	4.516	5.197
	(7.24)	(5.05)	(27.10)	(20.18)
Log of wealth	0.583	0.650	2.801	3.257
	(6.96)	(6.08)	(57.68)	(40.25)
Work	0.124	2.117	0.214	3.220
	(0.46)	(2.34)	(1.26)	(4.44)
Sigma u	56.430	60.184	41.376	49.231
	(99.43)	(89.07)	(115.70)	(96.50)
Sigma e	47.670	54.862	41.990	52.100
	(159.15)	(153.66)	(230.37)	(188.00)
N		71,108**		70,717**

** Time dummies and intercept included but not reported.*

*** Excluding item nonresponse*

Table 11 Covariance Matrices of individual effects: Complete Model

		Probability		Rounding		Nonresponse		50/50	
Worklm	**Probability**	365.87	*						
	Rounding	6.51	*	0.18	*				
	Nonresponse	1.98	*	0.28	*	2.29	*		
	50/50	8.50	*	0.40	*	0.85	*	1.39	*
Work62	**Probability**	1958.11	*						
	Rounding	0.71	*	0.39	*				
	Nonresponse	3.94	*	-0.03	*	0.49	*		
	50/50	9.14	*	0.08	*	0.79	*	2.29	*
Inher	**Probability**	3184.35	*						
	Rounding	26.44	*	1.11	*				
	Nonresponse	9.23	*	0.11	*	1.91	*		
	50/50	21.27	*	0.09	*	2.00	*	2.40	*
Beq10	**Probability**	1711.96	*						
	Rounding	-3.69	*	0.44	*				
	Nonresponse	-15.29	*	0.23	*	2.11	*		
	50/50	-117.67	*	0.85	*	8.68	*	38.12	*